Single-cell time courses of mRNA transport and translation kinetics

Anita Reiser

München 2019

Einzelzell-Zeitverläufe von mRNA Transport und Translationskinetiken

Anita Reiser

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Anita Reiser
aus Günzburg

München, den 14.05.2019

Erstgutachter: Prof. Dr. Joachim O. Rädler
Zweitgutachter: Prof. Dr. Hermann E. Gaub
Tag der mündlichen Prüfung: 18.07.2019

Single-cell time courses of mRNA transport and translation kinetics

Anita Reiser

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Anita Reiser
aus Günzburg

München, den 14.05.2019

Erstgutachter: Prof. Dr. Joachim O. Rädler
Zweitgutachter: Prof. Dr. Hermann E. Gaub
Tag der mündlichen Prüfung: 18.07.2019

Contents

Zusammenfassung

Die Messenger RNA (mRNA) spielt eine zentrale Rolle in der Genregulation. Diese wird im Zellkern durch das Ablesen eines Gens hergestellt und danach im Zellinneren durch Ribosome in Proteine übersetzt. Die mRNA ist aufgrund ihrer transienten Natur ein sehr flexibles Molekül, da sie weder ein permanenter genetischer Informationsspeicher ist noch ein funktionelles Produkt darstellt. Mit Hilfe von Nanofähren kann synthetische mRNA in Zellen transferiert und die Translation, d.h. die zell-interne Produktion beliebiger Proteine, erreicht werden. Dabei ist sowohl der mRNA Transportprozess in Zellen als auch die Translationskinetik wenig verstanden. Dies hat jedoch große Bedeutung für den therapeutischen Einsatz von mRNAs in der Nanomedizin.

In dieser Arbeit wurde ein *in vitro* Assay entwickelt, der es ermöglicht, individuelle Einzelzellkinetiken innerhalb einer Zellpopulation sowohl bezüglich des mRNA Nanopartikeltransports, als auch der Translationskinetiken zu messen. Der Assay beruht auf Zeitraffermikroskopie von Zellen auf proteinbeschichteten, mikrostrukturierten Feldern, sogenannten Einzelzell-Arrays, zur parallelisierten Aufzeichnung von Einzelzell-Zeitverläufen der Expressionskinetik eines Fluoreszenzproteins. Die Proteinexpression wird dabei durch den Transport von synthetischer mRNA ins Zellinnere ausgelöst. Das Array wurde mit einem Perfusionssystem verbunden, um einen Flüssigkeitsaustausch während der Beobachtung am Mikroskop zu ermöglichen. Dadurch war erstmals die Aufnahme der gesamten Fluoreszenzprotein-Expressionskinetik einschließlich der Anfangsphase in Abhängigkeit eines definierten Zeitintervalls der mRNA Nanopartikelexposition messbar. Die Translationskinetik wurde dabei durch das Auslesen der Fluoreszenzintensitätszeitverläufe einzelner Zellen auf der Mikrostrukturierung erhalten.

Um die Zeitverläufe der Translation zu quantifizieren, wurden in dieser Arbeit verschiedene biochemische Raten-Modelle auf der Grundlage gewöhnlicher Differentialgleichungen untersucht. Das Modell, das die experimentellen Daten am besten beschreibt, wurde genutzt, um die kinetischen Raten jeder einzelnen Zelle zu ermitteln. Die Einzelzellparameter ermöglichten dabei die Bestimmung der Populationsverteilung jedes Parameters und damit der Zell-zu-Zell Variabilität.

Im zweiten Teil der Arbeit wurde die präzise Messung von Einzelzellkinetiken genutzt, um den Expressionsbeginn jeder Zelle zu bestimmen. Dieser gibt Auskunft über die Dauer des mRNA Transportprozesses. Wir zeigten anhand von zwei Modellsystemen, Lipofectamine und Lipidnanopartikeln (LNP), dass die Dauer des mRNA Transports nicht mit der Effizienz des Transports in einzelnen Zellen korreliert. Wir konnten zeigen, dass sich diese Parameter in Abhängigkeit steigender Serumproteinkonzentration systematisch ändern. Die Serumproteinadsorption beeinflusst sowohl den Zeitpunkt als auch die Effizienz von Lipofectamine negativ, während sie im Gegensatz dazu den Zeitpunkt und die Effizienz der LNP-vermittelten Transfektion verbessert. Dies spielt für die medizinische Anwen-

dung eine große Rolle da die Nanopartikel üblicherweise über das Blut verabreicht werden und dabei in Kontakt mit Serumproteinen kommen.

Der entwickelte Assay stellt zusätzlich eine neue Chance zur Quantifizierung der mRNA Halbwertszeit in lebenden Zellen dar. Die mRNA Halbwertszeit ist ein Schlüsselparameter der Translationsregulation und unter Anderem abhängig von der poly(A) Schwanzlänge. Es ist jedoch nicht quantifiziert, wie genau die mRNA Halbwertszeit von der poly(A) Schwanzlänge abhängt. Zu diesem Zweck wurden mRNA Konstrukte mit zehn unterschiedlichen poly(A) Längen untersucht. Auf Grundlage der Messergebnisse formulierten wir erstmals ein einfaches Modell, das die Beziehung zwischen poly(A) Länge und mRNA Lebenszeit mittels eines Potenzgestezes mit einem Exponenten von 0,53 beschreibt.

Nutzt man für den Einzelzell-Translationsassay die Aufnahme eines weiteren Fluoreszenzsignals sind zusätzliche intrazelluläre Korrelationsstudien möglichen. Dies kann beispielsweise eingesetzt werden um die mRNA-Nanopartikel detektierbar zu machen wodurch der zeitliche Abstand zwischen Nanopartikeladsorption an der Zelloberfläche und dem Proteinexpressionsbeginn messbar ist. Durch eine degradationsabhängige Fluoreszenzmarkierung der mRNA könnten zukünftig mit Hilfe der Einzelzellarrays die zeitlichen Veränderungen der mRNA- und Proteinkonzentration zur genaueren Auflösung der mRNA Degradation simultan gemessen werden.

Summary

Messenger RNA (mRNA) plays a central role in gene regulation. The mRNA is produced in the cell nucleus by transcribing a gene and is afterwards translated into proteins by ribosomes in the cytoplasm. Due to its transient nature, mRNA is a very flexible molecule because it is neither a permanent storage for genetic information nor a functional product. With the help of nanocarriers, synthetic mRNA can be transferred into cells and the translation, the cell-internal production of any protein, can be achieved. Both the mRNA transport process in cells and the translation kinetics are poorly understood. However, these are of great importance for the therapeutic use of mRNAs in nanomedicine.

In this thesis, an in vitro assay was developed which enables the measurement of individual single-cell kinetics within a cell population with regard to mRNA nanoparticle transport as well as translation kinetics. The assay is based on time-lapse microscopy of cells on protein-coated microstructured fields, so-called single-cell arrays, for parallel recording of single-cell time courses of the expression kinetics of a fluorescent protein. Protein expression is triggered by the transport of synthetic mRNA into the cell interior. The array was connected to a perfusion system in order to enable the exchange of fluids during observation at the microscope. This made it possible to measure the entire fluorescence protein expression kinetics including the initial phase as a function of a defined time interval of mRNA nanoparticle exposure for the first time. The translation kinetics were obtained by reading out the fluorescence intensity time courses of individual cells on the microstructured fields.

In order to quantify the time courses of the translation, different biochemical rate models based on ordinary differential equations were investigated. The model that described the experimental data best was used to determine the kinetic rates per cell. The single cell parameters allowed the determination of the population distribution of each parameter and thus the cell-to-cell variability.

In the second part of the study, the precise measurement of single-cell kinetics was used to determine the onset of expression of each cell. This provides information on the duration of the mRNA transport process. Using two model systems, lipofectamine and lipid nanoparticles (LNP), we showed that the duration of mRNA transport does not correlate with the efficiency of transport in individual cells. We were able to show that these parameters change systematically with increasing serum protein concentration. Serum protein adsorption negatively affects both the timing and the efficiency of lipofectamine, whereas it improves the timing and efficiency of LNP-mediated transfection. This is an important finding, as nanoparticles are usually administered via the blood for therapeutic purposes and thus come into contact with serum proteins.

The developed assay also represents a new opportunity to quantify the mRNA half-life in

living cells. The mRNA half-life is a key parameter of translation regulation and depends, among other things, on the poly(A) tail length. However, how exactly the mRNA half-life depends on the poly(A) tail length has not been quantified. To this end, mRNA constructs with ten different poly(A) lengths were investigated. Based on the measurement results, we formulated a simple model that describes the relationship between poly(A) length and mRNA half-life using a power law function with an exponent of 0.53.

If a further fluorescence signal is recorded for the single cell translation assay, additional intracellular correlation studies are possible. This can be used, for instance, to make the mRNA nanoparticles detectable, which makes it possible to measure the time interval between nanoparticle adsorption on the cell surface and the onset of protein expression. Via a degradation-dependent fluorescence labeling of the mRNA, the temporal changes of the mRNA and protein concentration could be measured simultaneously, which would allow for a more accurate resolution of the mRNA degradation with help of the single cell arrays.

Associated publications and manuscripts

[P1] **Reiser, A.**; Zorn, M. L.; Murschhauser, A. & Rädler, J. O. Ertl, P. & Rothbauer, M. (Eds.) Chapter 4: Single Cell Microarrays Fabricated by Microscale Plasma-Initiated Protein Patterning (µPIPP) in Cell-Based Microarrays: Methods and Protocols, Springer New York, 2018, 41-54

[M1] **Reiser, A.**; Gruber, S.; Woschée, D.; Rädler, J. O. Cell confinement effects on the signal-to-noise ratio in single-cell fluorescence time-lapse studies, to be submitted

[P2] Fröhlich, F.*; **Reiser, A.***; Fink, L.; Woschee, D.; Ligon, T.; Theis, F.; Rädler, J. O. & Hasenauer, J. Multi-Experiment Nonlinear Mixed Effect Modeling of Single-Cell Translation Kinetics after Transfection, npj Systems Biology and Application, 2018

[P3] **Reiser, A.**; Woschée, D.; Mehrotra, N.; Krzysztoń, R.; Strey, H. H.; Rädler, J. O. Correlation of mRNA delivery timing and protein expression in lipid-based transfection, bioRxiv, 2019, submitted to Integrative Biology

[M2] Dabkowska, A.; Frank, K.; **Reiser, A.**; Krzysztoń, R.; Yabez Artet, M.; Nickel, B.; Rädler, J. O.; Lindfors, L. The pH dependent phase transition in lipid nanoparticle cores leads to changes of protein expression in single cells, to be submitted

[M3] **Reiser, A.**; Enzedam, S.; Trepotec, Z.; Woschée, D.; Aneja, M.; Rudolph, C.; Plank, C.; Rädler, J. O. Single-cell studies of functional mRNA half-life distributions depending on poly(A) tail length, to be submitted

[M4] Krzysztoń, R.; Woschée, D.; **Reiser, A.**; Schwake, G.; Strey, H. H.; Rädler, J. O. Single-Cell Kinetics of siRNA-Mediated mRNA Degradation, accepted by Nanomedicine, 2019

* authors contributed equally to the work

1. Introduction

"Biology is the most powerful technology ever created. DNA is software, proteins are hardware, cells are factories."

Arvind Gupta

The cell factory from a physical perspective

Cells are the smallest units of life and probably the most essential process for keeping a cell alive is the expression of proteins. Even Erwin Schrödinger already dealt with the question "What is life?" and tried to answer it by explaining living matter with physical principles [1]. From a physical point of view, the living cell is an open system out of equilibrium [2]. Cells constantly consume energy in order to build up and dismantle new components and thus are in constant exchange with their environment. Consequently, the formation and degradation of proteins is critical, as their abundance is controlled by gene regulatory networks. A simplification of these complex networks is summarized in the central dogma of molecular biology: DNA is transcribed into messenger RNA (mRNA) followed by translation of the mRNA into protein [3].

In order to find general laws for these complex networks, the field of systems biology emerged [4]. Since living cells are subject to evolution, design principles have evolved that enable the formulation of mathematical frameworks in order to describe biology [5]. Especially for gene regulation, a large number of models were developed that are based on building block circuits which are depend on components with naturally occurring fluctuations and thermal noise [6, 7, 8, 9]. These networks need to function over a large temporal and spatial range in order to sustain life. As an example, the typical size of a mammalian cell is 10,000 μm^3 and the typical diameter of a protein is in the range of $1 \cdot 10^{-3}$ μm^3. It takes a protein with a diffusion rate of 10 $\mu m^2/s$ 100 s to cross the cell which on its own already has a large influence on reaction kinetics for proteins of low concentration. Furthermore, the human genome contains ~30,000 genes that encode for proteins with varying abundance and half-life. Taking transcription and translation together, it takes roughly 60 min to produce a protein. In order to keep the machinery running, a cell contains 10^8 ribosomes that translate mRNA with a rate of 5 amino acids/s into proteins (numbers are taken from [5, 10]). All these properties form the gene regulatory network. Even if each of the network components is fully understood, the complex interacting system has emergent properties. In order to understand the collaborative functioning of at least parts of the network, a system-level perspective is needed to quantitatively analyze it as a function of time [8]. If one of the proteins in the network is no longer produced

correctly, the system may switch from a healthy to a disease state. The error leading to the system's change can occur on each level of gene regulation.

Targeted alteration of protein expression kinetics

In order to correct for errors within the gene regulatory network, the research field of gene therapy has become more prominent during the last decades to enable the targeted manipulation of affected protein expression. Gene therapy can be understood as the delivery of protein percursors or molecules regulating protein expression. The basic idea is to introduce missing, replace deficient, or regulate the expression of genetic information [11]. These functions can be controlled by the delivery of DNA, mRNA, or small interfering RNA (siRNA) among other molecules. In 2018, the first siRNA based drug was approved by the US Food and Drug Administration, which is a key event for the field and was therefore taking into account as the breakthrough of the year by the journal Science [12, 13]. As the artificial delivery of DNA has the risk of permanently changing the genome, the research for gene therapy focuses on RNA based approaches. Beside siRNA, mRNA has become a potential new drug class as it is the middleman of the central dogma [14, 15]. mRNA itself is a highly flexible molecule due to its transient nature as it is neither a permanent storage of genetic information nor a functional product [15]. A targeted alteration of protein expression kinetics by introducing artificial mRNA is only possible if the first obstacle, the cell membrane, is overcome [16]. As mRNA is a large, negatively charged molecule, the event of spontaneous crossing of the cell membrane is very rare due to the electrostatic repulsion of the anionic cell membrane [14, 17]. Furthermore, the mRNA molecules are inherently unstable, as they are constantly degraded by omnipresent nucleases which lead to a limited half-life and thus to a restricted duration of protein expression [15].

In order to transport and protect mRNA molecules, they need to be complexed in nanoparticles that enable mRNA release into the cytosol to trigger protein expression. For the purpose of nanoparticle formation that encapsulate mRNA, lipid-based delivery systems are one of the most often used systems [18]. The mRNA nanoparticle transport depends on the formation and structural changes of these nanoparticles during their interaction with the cell's surface and inside the cell. Thereby, the mRNA nanoparticle formation has a soft matter nature, as mRNA, which is a bioploymer, interacts with lipids to assemble to higher structures. Furthermore, these nanoparticles have to interact with the cell's lipid bilayer membranes, which are complex systems on their own [19], for successful release of the mRNA cargo into the cytosol. It is not fully understood what the transport kinetics execute and how physico-chemical changes of the environment influence the mRNA nanoparticle transport.

After the release of mRNA from the nanoparticle into the cytosol, the mRNA triggers the expression of the encoded protein and thus alters the regulatory network. The process of mRNA delivery followed by protein translation can be described as multi-step kinetics that can be predicted by mathematical models [20, 21]. Quantitative description of these processes are necessary due to the naturally occurring molecular and phenotypic

differences even within an isogenic cell population and the additionally arising variance of delivered mRNA molecules [6, 22]. The inherent heterogeneity of cell populations is a fundamental property of cellular systems and the measurement of ensemble averages to determine the average cell response can mask the information contained in the cell-to-cell variability [22, 23]. In order to identify the heterogeneity of a cell population, techniques with single-cell resolution are needed. In the case of gene expression, the process depends on a relatively small number of molecules which lead to a high variance within a cell population and especially across different cell types and cell states [24, 25]. The same is the case for the artificial expression of mRNAs after nanoparticle transport. Therefore, the protein expression kinetic of a single cell is not only dependent on the copy number of the delivered mRNAs, but also on fluctuations regarding the rates that define the biochemical reaction of translation e.g. the mRNA half-life. The development of new single-cell techniques that enable the observation of kinetics or the determination of mRNA abundance in thousands of cells at one time point was acknowledged as the breakthrough of last year in 2018 [26].

Quantification of translation after mRNA transfection

A large number of methods exist that enable the quantification of mRNA or protein abundance to certain time points and allow for their correlation [27, 28, 29]. These methods are often limited to bulk measurements and it is often not possible to measure the abundances of the same cells at different time points which results in a low time resolution and aggravate kinetic observation. Even single-cell measurements that investigate a large number of genes lack good time resolution relevant for the process of interest [30]. In order to monitor the kinetics of single cells, the development of assays based on time-lapse measurements increased. The use of quantitative time-lapse experiments allow the recording of spatio-temporal dynamics of thousands of cells in parallel [31]. The observation of fluorescence intensity time courses of reporter proteins proved to be a powerful tool to investigate gene expression at the single-cell level [32, 33]. Many of these assays rely on the use of microfluidic chips or microstructured arrays for parallel observations and control of extracellular factors [34, 35, 36]. The quantification of fluorescence reporter protein time courses after mRNA nanoparticle transport in single-cells is possible using live imaging of single-cell arrays (LISCA) and results in the record of hundreds of single-cell protein expression time courses per experiment [33]. An advantage of the assay is that it is very flexible, as many parameters are controllable and can be systematically varied for quantitative studies. Several different scientific aspects can be addressed by the variation of the used nanoparticle system e.g. the chemical composition or by mRNA sequence modifications, which lead to varied protein expression kinetics.

The outline of this thesis is as follows: First, an overview of the fundamental concepts are given in chapter 2. The basics of artificial gene expression is explained with special emphasis on fluorescent reporter protein expression after mRNA transfection. In order to quantify protein translation kinetics after transfection, system biological approaches were used which are based on mathematical modeling of biochemical rate equations. In order

to obtain reporter protein expression time courses of living cells over a long time period, a single-cell translation assay based on a previously published method [37] was improved and the main elements of the assay are discussed. In chapter 3, potential sources for noise within the single-cell time courses are discussed. The noise mainly hinges on the ability of cell confinement, which varies dependent on the microstructure fabrication technique, and on the image processing algorithm to readout the time courses of the reporter fluorescence intensity per cell.

The use of single-cell time-lapse experiments improved the quantitative understanding of cellular processes and their heterogeneity. However, the estimated parameters that describe the observed processes depend on the model and the estimation does not necessarily provide robust results for experimental replicates. In chapter 4, a newly developed modeling approach is discussed for the fitting of mechanistic pathway models to single-cell protein expression time courses. It was shown that the integration of perturbation experiments allowed for the robust estimation of the cell-to-cell variability of cells transfected with mRNA across experimental replicates.

As discussed previously, the use of mRNA to investigate translation kinetics relies on the successful transport of mRNA nanoparticles. Yet, it is not well understood how the transport process depends on the choice of nanoparticle system or extracellular properties. In chapter 5, a translation-maturation model based on ordinary differential equations (ODEs) is introduced. By fitting this model to hundreds of protein expression time courses, the onset time point of expression as well as the expression rate are determined. It was shown that the timing of delivery is not correlated with expression efficiency on the single-cell level autonomous of the investigated model systems used as nanoparticle system.

Beside the mRNA nanoparticle transport, a more quantitative understanding of the half-life regulation of transfected mRNA is needed to enable targeted alteration of translation. The half-life of mRNAs is known to be based on various sequence elements like the length of the poly(A) tail, a stretch of adenosine residues at the 3' end of the mRNA [38]. However, there is no quantitative description how changes in poly(A) tail length change the half-life of mRNAs and thus affect the protein expression kinetics at the single-cell level. In chapter 6, mRNA constructs of ten varied poly(A) tail lengths, which covered the range of naturally occurring poly(A) tail lengths, were observed by LISCA experiments and analyzed using the translation-maturation model. It was shown that the only parameter distribution that systematically shifted with increasing poly(A) tail length was the mRNA half-life. Based on the half-life distributions, a simple model was found that described the dependency of mRNA half-life on poly(A) tail length.

Additional information is obtained by introducing a second fluorescent marker for each cell as described in chapter 7. The dual color approach allows, for instance, correlation studies of event times or kinetic rates of different dynamics. Finally, the results and gained insights of the described studies are summarized in chapter 8. The appendix of the thesis contains a brief description of the used experimental protocols (Appendix A) and the investigated pathway models for translation after mRNA delivery (Appendix B).

2. Fundamental concepts and methods

The aim of this thesis is to investigate single-cell translation kinetics after the introduction of artificial mRNA to quantify the cell-to-cell variability regarding the various scientific questions discussed in chapter 3 to chapter 7. Helpful background information and an explanation of the experimental assay used in this thesis is given in the following sections.

2.1. Artificial gene expression

The central dogma of molecular biology

Most processes in living cells rely on functioning proteins. These are produced and regulated by gene expression, which is described by the central dogma of molecular biology [3]. Gene expression takes place in a two-step process: The transcription of DNA to mRNA followed by translation of protein from the transcribed mRNA. The production of mRNA and protein as well as the degradation of both molecules are characterized by kinetic rates. The regulation of these rates offers control of mRNA and protein levels defining the cellular states. The systematic quantification of mRNA and protein concentrations and the respective degradation rates allow for global characterization of gene expression in mammalian cells [27, 39, 40]. Schwanhäusser et al. analyzed >5,000 genes in mouse fibroblasts and observed median rates for transcription of two mRNA molecules per hour and a translation rate of 40 proteins per mRNA molecule and hour with respective mRNA half-lives of 7.6-9 h and 46 h for proteins (Figure 2.1.) [27]. However, rates are varying over several orders of magnitude for different genes. This high variability is crucial in living systems as proteins need to fulfill a huge variety of functions and protein levels need to be flexibly adjusted to changing conditions. In order to understand how cells behave under different conditions and how they regulate and adapt their protein household is of major interest. It is an ongoing debate if transcription or translation is more important for gene regulation [41, 42, 43]. A study by Vogel et al. showed that the variations of protein abundance can be explained to 27% by transcriptional regulation and to 40% by processes related to translation and protein degradation (Figure 2.1.) [28]. Further studies confirm that post-transcriptional regulation is at least as important as transcriptional regulation for controlling protein levels in living cells [42, 44]. On the one hand, the studies quantify the dynamical range of gene regulation for a large number of genes. On the other hand, they do not take into account that there is not only a high variation between different genes but that cells of an isogenic population already have a high cell-to-cell variability [6, 45]. In order to investigate the population heterogeneity of the kinetic rates single-cell techniques are pivotal.

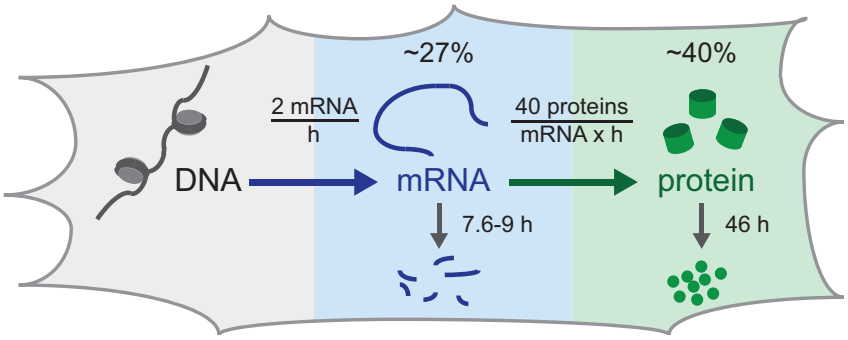

Figure 2.1.: Gene expression involves transcription of DNA to mRNA and translation of proteins from mRNA. The protein abundance can be controlled via the regulation of the kinetic rates for transcription or translation represented by median values measured in a global study for >5,000 genes [43]. About 30% of the protein concentration changes can thereby be explained by transcription regulation and 40% by translation regulation [42].

Transfection of nucleic acids

In order to understand the cell-to-cell variability of translation kinetics in a cell population there is need to quantify protein expression. For this purpose, reporter proteins are commonly used. To induce reporter protein expression in living cells it is necessary to introduce artificial molecules (plamsid DNA (pDNA) or mRNA), which carry the genetic code of the reporter protein. The process of nucleic acid (NA) delivery in cells is called transfection. This approach is commonly used to study reporter protein expression in cells and various transfection-methods were developed [15]. The introduction of mRNA has the advantage that the molecule only needs to enter the cytosol for protein expression, whereas pDNA has to be additionally introduced into the nucleus and posing the risk of permanent integration into the genome. A further advantage of mRNA introduction is that the observed protein expression is independent of transcriptional regulation, simplifying the regulation network. mRNA molecules commonly exists of a 5' cap, a 5' untranslated region (UTR), an open reading frame (ORF) coding for the protein, a 3' UTR, and a poly(A) tail at the 3' end. The minimal functional structure of mRNA consists of the 5' cap, an ORF, and a poly(A) tail. The stability of artificial mRNAs, which are *in vitro* transcribed (IVT), is influenced by the use of chemically modified nucleotides, the cap structure, the UTR sequences as well as the poly(A) tail lengths [46, 47].

In the framework of this thesis the transfection of mRNA was used to quantify translation kinetics. For this reason, the principles of mRNA transfection are described. The transfected mRNA needs to be transported inside the cells and translated for successful reporter protein expression (Figure 2.2.). For the transport, the mRNA needs to pass the cell membrane. As mRNAs have a typical molecular weight of 10^5 to 10^6 Da, they

are to large to simply diffuse into the cell. In addition, the negative charge of mRNAs and the anionic cell membrane lead to electrostatic repulsion. Therefor, the uptake of naked mRNA is very unlikely (less than 1 in 10,000 molecules) [15]. Consequently, mRNA molecules need to be encapsulated to enable cell membrane crossing. In general, the mRNA can be delivered by viral or non-viral vectors. Viral vectors typically have a very high efficiency but a higher risk compared to non-viral vectors as they often lead to immunological response of the targeted cells [48]. Viral free alternatives for transfection are for example microinjection, electroporation, or lipofection [46]. The term lipofection refers to the transfection of NA using lipid-based systems. For this method, a mixture of cationic lipids with neutral helper lipids is used to encapsulate the NA. Cationic lipids interact with the negatively charged NA leading to a complex formation often referred as lipoplex, which can be taken up by cells via endocytosis. Endocytosis is a cellular process that leads to an uptake of particles and fluid from the cellular environment by local invagination of the cell membrane. The invagination leads to the formation of new cell organelles, the endosomes, from where the uptaken particles can be released into the cytosol. A more detailed overview of lipid-based vectors and the endocytotic uptake is given in section 5.1. After the release of mRNA into the cytosol, it can be directly translated to proteins. For the observation of translation kinetics, reporter proteins based on luminescence or fluorescence are often used.

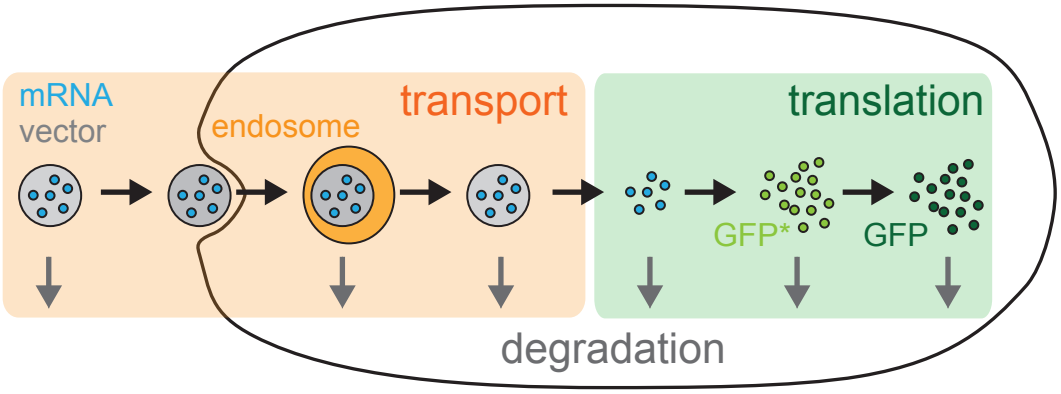

Figure 2.2.: The expression of artificial mRNA consists of two major processes: the nanocarrier transport and the mRNA translation. Both consist of further intermediate steps. The mRNA encapsulated in some vector is taken up by the cell via endocytosis and the released mRNA in the cytosol is translated into protein. At all intermediate steps degradation processes can take place.

Fluorescence reporter proteins

Fluorescence reporter proteins are widely used in biological applications to monitor the spatial and temporal changes of the expressed protein. One common reporter protein is the green fluorescent protein (GFP), originally derived from the jellyfish *Aequorea victoria* [49]. The GFP has a barrel structure consisting of eleven β-sheets with an α-helix inside

holding the chromophore [50]. In order to obtain a functional chromophore, the polypeptide backbone has to undergo a chemical transition called maturation involving folding, cyclisation, oxidation, and dehydration [51]. The three amino acids involved in maturation are Serin[65], Tyrosin[66], and Glycine[67] resulting in the fluorophor 4-hydroxybenzlidene imidazolinone (HBI) [52]. The maturation time itself depends on different environmental conditions and can vary from a few minutes up to an hour [51]. Depending on the observed kinetics it is therefore important to take the maturation step into account. GFP can be expressed in mammalian cells for live-cell imaging and is the basis for many other artificial proteins with engineered properties such as emission properties, or higher stability like red fluorescent protein or enhanced GFP (eGFP) [53].

2.2. Mathematical modeling of single-cell dynamics

Living cells of an isogenic population, which are genetically identical, show molecular and phenotypic differences in gene expression at the single-cell level [6, 22]. This heterogeneity can arise from stochastic noise [6, 7], varied cell-cycle states [54], or spatio-temporal differences of the environment [55]. Bulk measurements of a cell population can mask the cell-to-cell variability within a population, which can lead to misinterpretation of the investigated biological system because the heterogeneity, dynamic behavior, as well as the cell fate cannot be monitored [56]. Methods like fluorescence activated cell sorting (FACS), single-cell RNA sequencing, or mass cytometry enable determination of a snapshot of the cell-to-cell variability of thousands of cells in parallel [27, 57, 58]. However, these techniques lack the ability to measure the dynamical response of individual cells. For this purpose, time-lapse measurements with single-cell resolution are a suitable tool [31, 32]. In order to improve quantification and high-throughput screening of single-cell dynamics, cell microarrays and microfluidic devices were implemented for time-lapse measurements to sort, confine cells locally and simplify single-cell trajectory readout of a reporter signal [31, 59].

These techniques lead to new insights regarding population structures, the development of cells, and intracellular processes like apoptosis or gene expression [54, 58, 60, 61, 62, 63]. To quantify mechanistic insights, deterministic models based on ordinary differential equations (ODEs) are often employed [4, 64]. The models themselves are derived from biological knowledge of the process. Here, the cell is treated as a reacting system that experiences a time dependent input-signal and generates a certain output behavior [65]. In order to understand the cell's behavior, reaction rate equations are commonly used to describe biochemical networks of cellular processes [5]. These biochemical rate equations are based on ODEs and denoted as $\sum(\theta)$ in Equation 2.1 [66]:

$$\sum(\theta) = \begin{cases} \frac{d}{dt}x(t,\theta) & = f(x(t,\theta),\theta), \quad x(0,\theta) = x_0(\theta) \\ y(t_i,\theta) & = h(x(t_i,\theta),\theta) \end{cases} \tag{2.1}$$

The concentration of molecules that are part of the network are represented by dynamical state variables x and the rates of the reactions are given by the parameters θ. The initial

conditions of the ODEs are defined as $x_0(\theta)$. Typically, only some of the system variables are determined experimentally and these are mostly measured at discrete time points t_i. With this in mind, the translation dynamics after transfection of pDNA or synthetic mRNA was quantified using ODE-based mathematical models in previous work [20, 21, 33, 67] as well as in this thesis. All deterministic models quantifying reporter protein translation dynamics after mRNA transfection with their respective ODEs used in this thesis are listed in Appendix B. All models describe the translation process in living cells taking different aspects of translation such as ribosome binding to the mRNA or fluorescent protein maturation into account. The model needs to be carefully selected, weighing accuracy of the described biological process vs. over-estimation with too many free parameters. In order to restrict the number of parameters the translation process can be simplified by reducing a multi-step reaction to an one-step reaction, ignoring processes or stochastic fluctuations [68]. The fitting of the model to the experimental data provides the estimation of biological important parameters like mRNA or protein degradation rates.

2.3. Single-cell translation assay

In order to observe single-cell translation kinetics of thousands of cells in parallel, a single-cell translation assay was used. This assay combines live cell imaging of single-cell arrays (LISCA) with mRNA transfection to measure the translation dynamics. The transfected mRNA encodes for a fluorescent reporter protein like eGFP. The basic principles of this setup were developed and published in previous studies [21, 33, 37] and the assay was further improved in the framework of this thesis.

The general workflow of a LISCA experiment consists of the preparation of a cell microarray followed by scanning time-lapse measurement of fluorescent reporter expression kinetics and corresponding image analysis like illustrated in Figure 2.3.. The use of a microarray enables advanced single-cell measurements due to protein coated adhesion sites with standardized boundary conditions for every cell (Figure 2.3. A). Moreover, the alignment of thousands of cells on a micropattern (Figure 2.3. B) simplifies the image analysis significantly. After self-organization of seeded cells on the microarray [69], a scanning time-lapse measurement is set up to observe the total micropatterned area over a duration of several hours up to 30 h (Figure 2.3. C). Finally, image analysis is performed to read out the single-cell intensity trajectories of a fluorescence reporter (Figure 2.3. D). The following is a description of crucial steps of the single-cell translation assay, detailed protocols of these steps are attached in Appendix A.

2.3.1. Cell microarrays

Many techniques were developed to fabricate cell microarrays to investigate cellular processes [71, 72]. Microarrays are used for a variety of scientific fields like migration studies, cell-cell interaction, or the measurement of intracellular signals [73, 74, 75]. Different microarray fabrication techniques can be applied for single-cell translation assays, which are explained in more detail in section 3.1. including a discussion of their advantages and

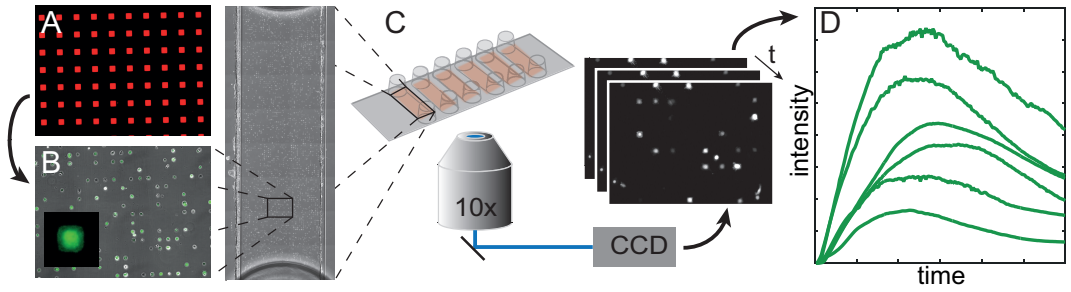

Figure 2.3.: LISCA is based on time-lapse microscopy of single-cell microarrays consisting of micropattern (A). The cells preferentially adhere on the protein coated squares yielding an array of cells aligned on the micropattern grid (B). The cells are treated in order to exhibit a fluorescently labeled reporter. The whole microstructured area is observed by scanning time-lapse microscopy (C). To obtain hundreds of single-cell traces, the recorded image stacks are background corrected and further processed to read out the fluorescence intensity over time (D). Reprinted with permission from [70].

drawbacks for translation studies. For most of the experiments the single-cell array was fabricated by microscale plasma induced protein patterning (µPIPP) [70]. The total array consists of six channels, as visulaized in Figure 2.3. C, with the same microstructure on the bottom of each channel. The cells align on the micropattern as illustrated in the phase contrast overview of one channel in Figure 2.4.. The alignment of the cells is even better visible for cells expressing eGFP after mRNA transfection like shown in Figure 2.4.. For this pattern geometry with a lattice spacing of 90 µm and a typical microstructured area of 4 mm x 10 mm roughly 5000 possible adhesion sites are available per channel.

For an experiment, the cells are seeded with a concentration of about $8{\cdot}10^4$ cells/ml, resulting in a high single-cell adhesion spot occupancy and minimizing the probability of double occupancy. The slide is incubated for at least 4 h at 37 °C before the assay with the time-lapse measurement is conducted (Figure A.2.). In order to improve occupancy, the cell culture medium can be exchanged after 1 h to remove non-adherent cells and to reduce the number of double occupied adhesion spots on the micropattern. Typically, HuH7 cells, a human hepato carcinoma cell line, were used for the presented transfection studies. The cell line was chosen because liver cells are an often targeted cell type for gene therapy and the HuH7 cells are a frequently used model system for liver cells [76, 77]. For other cell types it might be necessary to adjust the pattern geometry regarding the adhesion area and the lattice spacing.

2.3.2. Fluorescence time-lapse microscopy

A motorized inverted microscope (Nikon Eclipse Ti-E) was used for wide-field fluorescence time-lapse microscopy. The microscope was equipped with a 10x objective lens and suitable filter cubes for the respective fluorescent dyes. Images were recorded using a cooled CCD or CMOS camera with a LED light source. A heating chamber was used for stable temperature condition of 37 °C during the experiment with the possibility to control CO_2

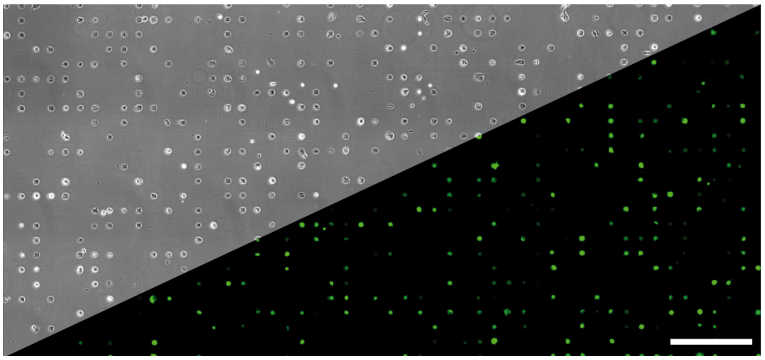

Figure 2.4.: Overview scan of a microstructured area of one single-cell array channel. The upper part shows the phase contrast image of HuH7 cells. The fluorescence image shows eGFP expressing cells after mRNA transfection on a square pattern. The scale bar corresponds to 500 µm.

level and humidity.

The scanning macro for time-lapse settings was controlled using NIS-Elements Advanced Research software (Nikon) to define important parameters like exposure times, position list, or time resolution. Time-lapse images were acquired with a frequency of 6 images per hour and a total duration of 10-30 h. For quantitative image analysis, it was important to keep imaging conditions as constant as possible to enable comparability within the measurement and to reduce technical noise between experimental replicates. For a typical single-cell assay 70-100 positions were scanned and an image stack for each position was recorded. The number of stacks mainly depends on the camera chip size, the total micropatterend area, and the time resolution of the experiment.

2.3.3. Transfection during time-lapse measurement

Perfusion system

In order to investigate the complete translation kinetics for each single cell in the experiment, it is crucial to perform mRNA transfection on the microscope during the time-lapse measurement. For this purpose, a tailored perfusion system for sterile liquid exchanges, as visulaized in Figure 2.5., is necessary. The perfusion system is made of PTFE microtubing with an inner diameter of 0.3 mm to keep the dead volume of the system low. For sterile injection of different fluids needle-free swabable valves are connected with female luer connectors to the microtubing. For connection of the tubing to the channel slide, inhouse made male luer teflon connectors were used. The perfusion system was cleaned by flushing ethanol after each experiment and autoclaving the system afterwards. Depending on the experimental design it is possible to wash each channel separately (Figure 2.5. A) or the series connection of two or more channels (Figure 2.5. B). Two channels are connected with a short piece of tube, so-called serial connector, with teflon connectors at both ends.

The perfusion system is connected to the channel slide after cell adhesion. It is impor-

tant for the experiment that the tubing is connected sterile and without air bubbles to avoid damaging the cells during perfusion. In order to connect the tubing, the cell culture medium L15, which is buffered CO_2 independently, supplemented with 10% FBS was used. This medium is used during the time-lapse measurement, excepted for the transfection. After connecting the perfusion system, the single-cell array was put in the pre-heated incubation box at 37 °C on the microscope. The time-lapse experiment was started after setting the position list and other microscope properties like exposure time or time resolution. A detailed description of how the perfusion system is assembled and used is given in Appendix A.

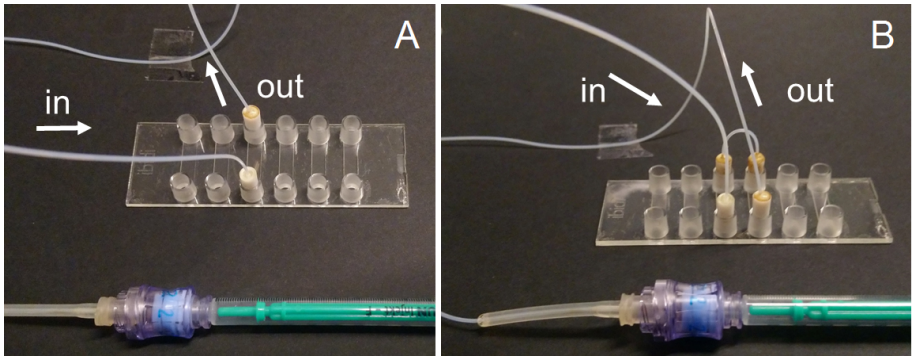

Figure 2.5.: Tailored perfusion system to enable liquid exchange during a time-lapse measurement using a syringe. (A) In order to flush one channel the inlet tube is pluged in one reservoir and the outlet tube is pluged in the opposite reservoir. (B) For the connection of two channels in a row, the inlet tube is connected to one reservoir. The opposite reservoir is connected with a neighboring one using a short serial connector tube. The respective opposite reservoir is connected with an outlet tube. The arrows indicate the flow direction.

mRNA lipoplex preparation

If not mentioned explicitly, the commercially available transfection agent Lipofectamine™ 2000 (L2000) was used to prepare mRNA lipoplexes. For lipoplex formation, mRNA was mixed with L2000 liposomes at a ratio of 2.5 µl L2000 per 1 µg mRNA. L2000 and mRNA were diluted separately in Opti-MEM™ and the L2000 dilution was incubated for 5 min at room temperature (RT). Afterwards, the L2000 solution was mixed with the mRNA solution at a ratio of 1:1 and incubated for further 20 min at RT with a mRNA concentration of 1 µg/ml. The mRNA lipoplex solution was diluted to the final mRNA concentration of 0.5 ng/µl. For transfection of one channel, 300 µl of the transfection solution are needed and for two channels 500 µl are sufficient to guarantee a complete fluid exchange to reach uniform transfection conditions in each of the microarray channels. Before adding the lipoplex solution, the channels were flushed with 37 °C warmed-up PBS. Afterwards, the lipoplexes were incubated for 1 h to enable lipoplex adsorption onto the cell surfaces and then washed with 37 °C warmed-up L15 medium supplemented with 10% FBS, which remains in the channels for the remaining measurement. All washing steps

were done on-stage during the time-lapse measurement and in between two time-lapse frames. If the break between two runs was not long enough for flushing all microarray channels, the measurement was paused and the resulting time shift was considered for data analysis.

2.3.4. Image stack processing

An image per time point is generated for each position yielding 70-100 image stacks per experiment. A representative time series of one position with eGFP expressing cells on a square micropattern is shown in Figure 2.6.. The transfection solution is added to the cells in between two frames of the time-lapse experiment and flushed out after the incubation period. The successfully transfected cells start expressing eGFP after a short delay time and the fluorescence signal intensity is increasing for several hours. Before the single-cell trajectories can be read out, the image stacks need to be corrected for background noise.

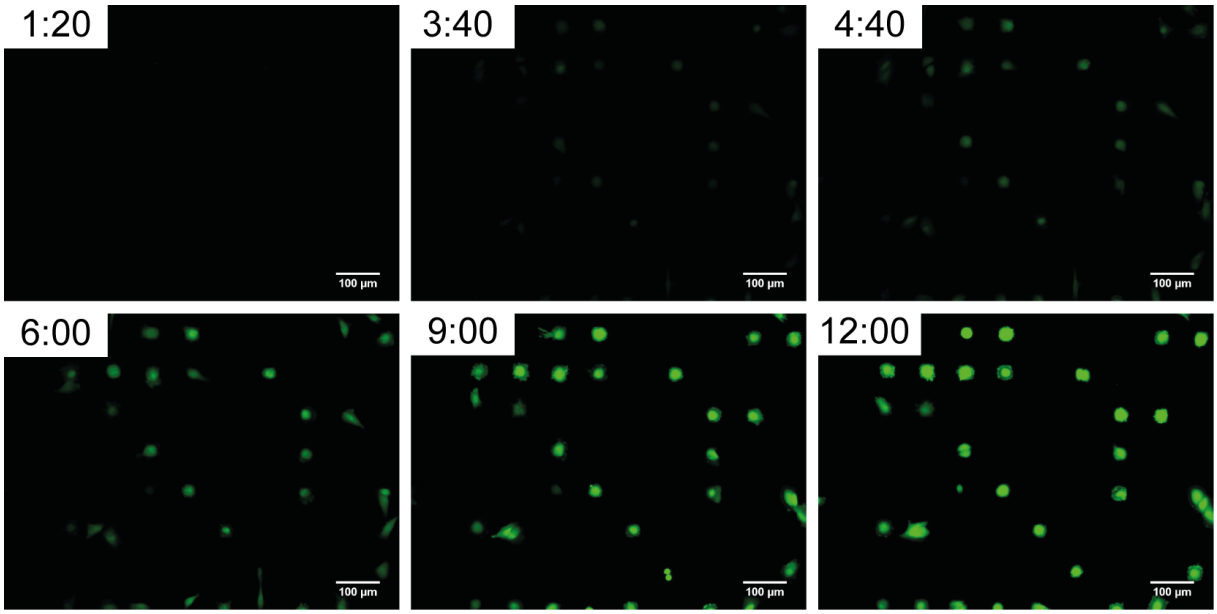

Figure 2.6.: Protein expression dynamic of HuH7 cells on a square micropattern transfected with eGFP mRNA. The transfection complexes were added at 1:20 h showing no detectable fluorescence signal as expected. The fluorescence of eGFP expressing cells can be detected by eye after around 3:40 h and the intensity is increasing with time. The scale bars correspond to 100 µm.

Background correction

The in-house background correction plug-in used in this thesis is based on a published algorithm [78] and implemented for the software ImageJ [79, 80].

The cellular fluorescence signal $s(x,t)$ is described by Equation 2.2:

$$s(x,t) = \frac{I(x,t) - b(t)g(x) - o(x)}{g(x)} \tag{2.2}$$

where $I(x,t)$ is the fluorescence image at the time point t and the pixel coordinate $x = (x_1, x_2)$, $b(t)$ is the background signal (that decreases over time due to photobleaching), $g(x)$ is the coordinate dependent illumination function, and $o(x)$ is a camera specific offset. In order to calculate the cellular signal each time-lapse image was subtracted by the illuminated background signal $B(x,t) = b(t)g(x) + o(x)$.

Each frame of an image stack is corrected individually by reconstruction of the background image by separating the cellular from the background signal using a density-based clustering approach (DBSCAN). For this purpose, each image is divided into small segments consisting of at least 56 pixel and the four statistical moments (mean, variance, skewness, and kurtosis) of the intensity distribution for each segment are calculated. The segments of the background form a very dense cluster if each segment is plotted in a four-dimensional space because the background is typically very similar within one frame with a weaker intensity compared to cellular signals. Each segment is assigned to either the background or cell cluster. Afterwards, the intensity of the pixels assigned to the background is set to the median value of the background segment intensity distribution and the cellular segments are set to zero. The background signal of the cellular segments is interpolated to reconstruct the complete background. The interpolated background image is afterwards subtracted from the original image to reconstruct the background-free cellular fluorescence signal $s(x,t)$ (Equation 2.2). The background corrected image stacks are ready for fluorescence trajectory readout.

Single-cell fluorescence time courses

In order to obtain single-cell fluorescence time courses, an in-house written semi-automated plug-in called Microwell Analysis (MWA) written by Christian Meggele was used. Based on the geometry of the micropattern, the orientation of the grid corresponding to the square micropattern is set manually for each image stack. The micropattern is marked by a grid and each outline of an adhesion spot is called well. The program calculates the mean intensity of each well for each frame of the image stack resulting in a mean fluorescence intensity trajectory over time. The wells occupied by cells need to be selected manually. Here, cells which double, die, or are not confined to the adhesion spot during the experiment need to be rejected to avoid bias of the data. The requirement is to only select wells that show a single-cell translation kinetics without any signal disturbances. As the exact cell contour is not taken into account by calculating the mean of the selected well, expansion or shrinking of cells lead to measurement noise. Representative trajectories of single-cell protein expression kinetics after mRNA transfection are shown in Figure 2.3. D. This way up to 3,000 single-cell trajectories are obtained per experiment and are the basis to determine the cell-to-cell variability of translation kinetics after mRNA transfection.

2.4. Single-cell parameter estimation

The fluorescence trajectories as measure for single-cell translation obtained by the MWA are further used for the parameter estimation. Therefore, a deterministic model that describes the translation process, is fitted to the data. The fluorescence intensity of each cell over time functions as model observable $y(t_i, \theta)$ that can be mapped to the variables $x(t_i, \theta)$ by a function h (Equation 2.1). A linear relation between the fluorescence intensity $y(t_i, \theta)$ and the protein concentration is assumed. In order to describe the relation between these two quantities, a scaling factor *scale* and a constant offset parameter *offset* are needed:

$$y(t_i, \theta) = scale \cdot h(x(t_i, \theta)) + offset \tag{2.3}$$

The scaling factor can be determined experimentally by calibrating the fluorescence intensity to the amount of protein using a microfluidic chip [21]. A spatial heterogeneity within the cells is neglected because the protein amount is measured as the average intensity signal of the whole cell and it is assumed that diffusion of the reporter protein within the cell is fast compared to the reaction rates. Furthermore, it is assumed that the experimental data $\tilde{y}(t_i)$ consists of the model observable $y(t_i, \theta)$ impaired by a measurement noise ϵ. The noise is assumed to be normally distributed $\epsilon(t_i) \sim \mathcal{N}(0, \sigma(t_i; \theta))$ with zero mean and an unknown standard deviation $\sigma(t_i; \theta)$, which is independent between different time points and can be estimated by:

$$\tilde{y}(t_i) = y(t_i, \theta) + \epsilon(t_i) \tag{2.4}$$

In order to estimate the single-cell parameters, a maximum likelihood estimation (MLE) was used [81]. The method is based on the assumption that the parameter values θ are deterministic but uncertain [82]. The parameter values of a given model are searched by maximizing the probability $p(D|\theta)$ to fit the experimental data $D = \{\tilde{y}(t_i)\}_{i=1}^{N}$. $p(D|\theta)$ is the probability for observing D when θ are the true parameters. The probability is called the likelihood function $L_D(\theta)$. Taking the given assumptions into consideration, the likelihood function is a distance measure between the experimental data $\tilde{y}(t_i)$ and the model observable $y(t_i, \theta)$ [79]:

$$L_D(\theta) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2(t_i; \theta)}} \cdot exp\left(-\frac{1}{2\sigma^2(t_i; \theta)} \cdot (\tilde{y}(t_i) - y(t_i, \theta))^2\right) \tag{2.5}$$

To estimate the parameter vector $\hat{\theta}$, the optimization problem is solved for each cell using a gradient-based algorithm [83]:

$$\hat{\theta} = argmax_\theta \left(-logL_D(\theta)\right) \tag{2.6}$$

For computational reasons it is more efficient to use the negative logarithm of the likelihood function. After selecting a suitable ODE based translation model, each experimental trajectory is fitted by the analytical or numerical solution of the model using the described MLE approach to determine the parameter values which yield the best fit of the data [84]. The population densities for each parameter can be estimated using kernel density estimations. More detailed information on the mathematical background of the model estimation and the implementation of the approach are given in [79, 80, 83].

3. Noise analysis of single-cell time courses on microstructures

For quantitative measurements of gene expression, a high number of cells is needed to reliably estimate the distribution, reflecting the heterogeneity within a population [85]. The use of micropatterend surfaces enables observation of thousands of cells in parallel over a long time period (typically a few hours, up to several days) under standardized conditions. Microarrays were used to investigate a variety of different phenomena e.g. morphology [71], migration [73, 86], or intracellular processes [33, 75]. Stable confinement of cells on microstructured areas is a crucial prerequisite for a reliable readout of the respective kinetics. In case of gene expression quantification, the readout of reporter fluorescence time courses was used to investigate the cell-to-cell variability [21, 32]. A powerful tool to study the protein expression dynamics in a highly parallel manner is the combination of single-cell microarrays with automated image processing [33, 83]. The reporter expression dynamics of each cell are recorded by measuring the repoter's fluorescence over time. Here, the measurement noise of these time courses is dependent on several factors like the micropatterning method or the image processing algorithm. Yet, it is not quantified how the quality of cell confinement on different micropattern influences the measurement noise and thus parameter estimation.

In this chapter, it is shown that the choice of micropattern technique and the related quality of cell confinement influences the noise in single-cell fluorescence time courses. Automated time-lapse fluorescence measurements with stably transfected cells expressing actin-eGFP on different microstructures were conducted. Based on the automated image analysis of the fluorescence time courses, noise in the time courses was characterized and quantified and subsequently related to the accuracy of cell confinement. Actin is a microfilament forming protein and is involved in cytoskeleton formation. Therefore, the fluorescent labeled actin is distributed over the whole cytoplasm of a cell. As actin-eGFP degradation and production should be in equilibrium, it is expected that an almost constant fluorescence intensity is observed over time. From a biological perspective, the single-cell mean fluorescence is expected to be almost noise free over time. The cell is a closed system regarding the fluorescence protein concentration as fluorescence proteins in the cytoplasm are too big to cross the cell membrane by diffusion. It follows that the noise in the fluorescence time courses must be due to other factors. The measurements showed that accurate cell confinement plays a pivotal role in limiting the noise of fluorescence single-cell time courses, which were analyzed by automated image analysis to a minimum level. Furthermore, the impact of different image processing algorithms is discussed. In addition to the analysis of stable protein expression, the d2eGFP expression after mRNA transfection was investigated. It was shown, that the type of microstructure not only

influences cell confinement but also the transfection efficiency. The transfection efficiency is determined as fraction of d2eGFP positive cells. A low transfection efficiency implies low statistics which is a serious disadvantage. This knowledge is valuable for any kind of quantitative analysis of single-cell fluorescence experiments like for the mRNA transfection studies discussed in the subsequent chapters and has the potential to reduce measurement noise in future cell experiments on cell dynamics. Parts of this chapter are based on the manuscript [M1] (listed in Associated publications and manuscripts) and the Bachelor's thesis of Sophia Gruber [87].

3.1. Single-cell microarray fabrication methods

The use of microstructered arrays has become an advantageous tool to systematically analyze cellular behavior [71, 88]. The ability to shape the microenvironment of cells and create standardized conditions by modulating the surface with molecules of different chemical properties is advantageous for quickly achieving large statistics. During the last years, many fabrication methods for microstructured surfaces were developed to investigate a variety of cellular phenomena like cell migration or signaling [72, 89]. In the scope of this work, three methods of micropatterning were tested and compared with respect to the quality of cell confinement.

3.1.1. Microscale plasma-initiated protein patterning µPIPP

The method of microscale plasma-initiated protein patterning is based on the fabrication of a 2-dimensional geometric pattern of locally varying surface properties, generated by plasma treatment like described in detail in [P1] (see Associated publications and manuscripts). Parts of the surface are activated by an oxygen plasma while the remaining area is protected by a 3-dimensional Poly(dimethylsiloxane) (PDMS) stamp during plasma treatment creating the prerequisite for a local variation in further treatment. A cell-repellent polymer electrostatically grafts the plasma activated areas. The non-activated areas are coated with an extracellular matrix protein like FN to generate cell adhesive regions. The final result are adhesive spots coated with protein and a cell-repellent interspace covered by polymer. The technique was used for a variety of studies including single-cell apoptosis or cytolysis assays [74, 75], migration studies [86] or mRNA transfection studies [33].

The important steps of the fabrication for a six-channel microarray using µPIPP are sketched in Figure 3.1.. Firstly, the PDMS stamps are formed using a silicon master of the desired geometry fabricated using photolithography. These stamps are cut to a proper size covering the area of one channel. Afterwards, the stamps are placed according to the channel positions on a coverslip (Figure 3.1. A) and treated with oxygen plasma to activate the unprotected areas (Figure 3.1. B). These areas turn hydrophilic due to the plasma exposure. They are passivated to avoid cell adhesion by adding a polymer solution of poly-L-lysine grafted poly(ethylene glycol) called PLL-g-PEG to the surface, which is

drawn under the stamps by capillary forces (Figure 3.1. C). The positively charged PLL-chains electrostatically interact with the hydrophilic surface and cover the whole plasma treated area. The PEG side chains of PLL-g-PEG render the coated surface protein and cell repellent. After washing out unbound PLL-g-PEG with destilled water, the stamps are removed and the coverslip is sticked to a six-channel slide (Figure 3.1. D). In order to coat the initially protected areas, a FN protein solution is added to each channel (Figure 3.1. E). After a final washing step to remove unbound FN, the single-cell array is ready to use (Figure 3.1. F). The treated surface results in micrometer-sized squares of cell adhesive FN adhesion spots surrounded by a passivated area of PLL-g-PEG. The cells preferentially adhere on the FN squares leading to a self-sorting of cells on the micropattern [69]. The micropattern used in this study exists of squares with an edge length of 30 µm coated with FN and a distance of 60 µm between neighboring squares.

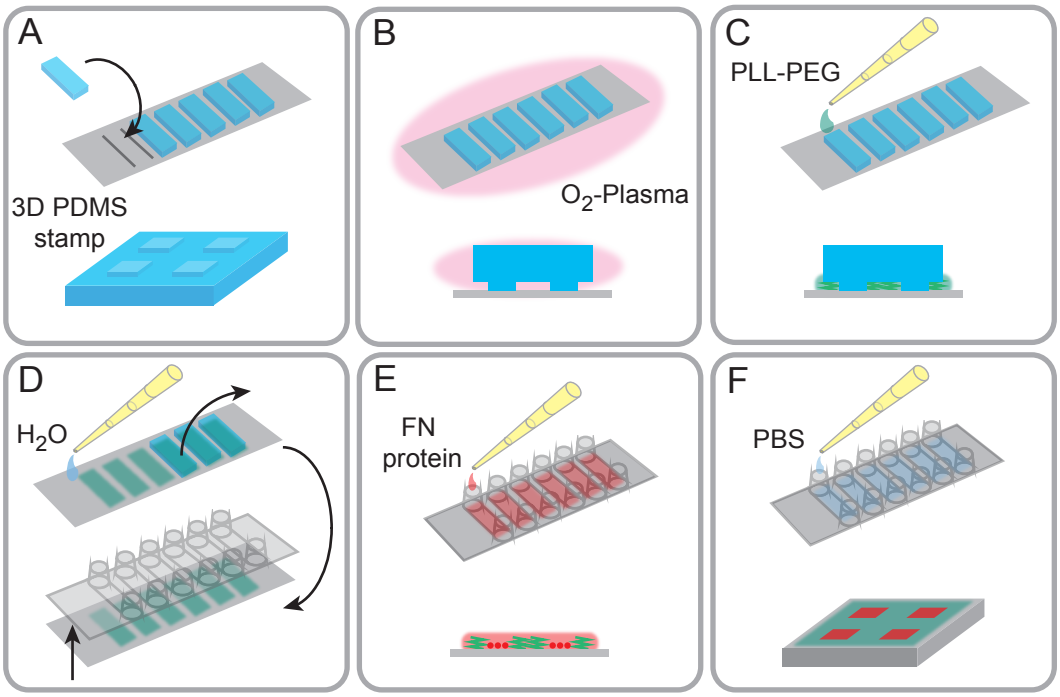

Figure 3.1.: Schematic overview of the fabrication of a µPIPP slide including: (A) stamp positioning, (B) oxygen plasma treatment, (C) passivation by adding a drop of PLL-PEG to every stamp, (D) washing unbound PLL-PEG away, removig the stamps, and sticking the six-channel slide to the coverslip, (E) FN protein coating, and (F) washing with PBS to remove unbound FN proteins. These steps lead to a 2D microarrray of FN protein (red squares) and a cell-repellent interface (green). Adapted and reprinted with permission from [70].

3.1.2. Microstructured gold pattern

The second investigated technique to fabricate a 2-dimensional micropattern is based on the same geometry as the µPIPP pattern but made of gold squares and a passivated

interspace coated with Pluronic F-127. Pluronic F-127 is a triblock polymer consisting of PEG-chains with a poly(phenylenoxid) inbetween the PEG chains [90].

The samples of gold micropattern were provided by the group of Philipp Paulitschke and the technique is described in the supporting information of [91]. A thin layer of photoresist is applied on a coverslip by spin coating. Afterwards, the photoresist is illuminated through a chrome mask with the desired square pattern resulting in a locally varying resist-solubility. The illuminated photoresist is removed resulting in a coverslip with square cavities. Then, to improve adhesion, a thin titanium layer (∼3 µm) is first evaporated on the coverslip followed by the evaporation of a gold layer (∼7 µm). In order to obtain the final gold pattern, the remaining photoresist is removed. The coverslip with the goldpattern is stuck to a six-channel slide and the hydrophobic interspace between the squares is grafted with Pluronic-F-127.

Cells are able to adhere on the gold surface in principle. But it was observed that the adhesion on the gold structures took longer than on the FN micropattern. The cell adhesion was improved by coating the gold islands with FN after passivation which led to a comparable adhesion behavior for cells on gold-FN structures and cells on µPIPP adhesion spots.

3.1.3. Micromolding in capillaries MIMIC

As a third method, 3-dimensional cavities were tested to confine the cells by physical borders instead of chemical based confinement by cell repellent polymers like for the 2-dimensional methods. An advantage of 3-dimensional micropattern is the higher adhesion spot density compared to 2-dimensional patterns. For 2-dimensional patterns the distance between neighboring spots needs to be large enough to minimize the probability that cells span over the passivated area to the next adhesion spot. This is not necessary for 3-dimensional pattern due to the physical boundary. The distance between cavities is mainly restricted by the structure stability. One method to produce cavities with a diameter suitable for single-cell studies is by micromolding in capillaries referred to as MIMIC. The method was developed to fabricate patterned structures on length scales of micrometers [92] and further improved for high-throughput single-cell assays [35, 93, 94].

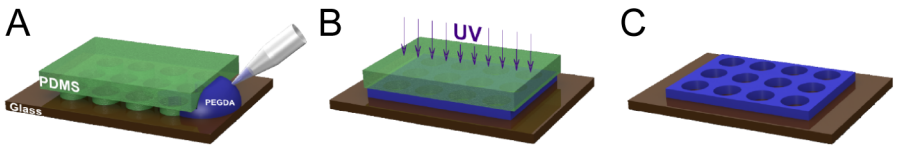

Figure 3.2.: Fabrication of µ-wells using the MIMIC method [95]: (A) A drop of PEGDA solution is pulled under the PDMS stamp by cappilary forces. (B) Due to UV-illumination the PEGDA is cured. (C) After removing the PDMS stamp the cured µ-wells are left on the substrate. Reprinted from [95].

Cylindrical cavities referred to as µ-wells were fabricated by using PDMS stamps with a pillar structure of 35 µm in diameter and 30 µm height for each pillar. The stamps were

placed on coverslips with a hydrophilic surface analog to the fabrication of µPIPP pattern and the surfaces were activated by argon plasma. A mixture of PEG-diacrylate (PEG-DA) with photoinitiator was added underneath the stamps by capillary forces (Figure 3.2. A). The PEG-DA solution was cured under UV-light in order to crosslink the PEG-DA (Figure 3.2. B) wherein the reaction is initiated by creation of reactive species due to radiation of the photoinitiator [96]. After UV treatment, the PDMS stamp was removed carefully and the µ-well structures remain on the coverslip (Figure 3.2. C). The coverslip was sticked to a six-channel slide as for the other two methods. Before the MIMIC array was ready to be used, the channels are washed with ethanol to avoid bacterial contamination. This was followed by a washing step with sterile water to remove non-crosslinked polymer as well as ethonal to avoid cytotoxicity effects.

3.2. Noise analysis of single-cell time courses

The experiments were performed similar to the description in section 2.3.. After the preparation of the respective microarrays, the cells were seeded 4 h before the time-lapse measurement. For the time-lapse experiment, phase contrast images and fluorescence images of the eGFP signal were recorded for each position with a 10 min time resolution for a duration of 20 h. The experiments with actin-eGFP expressing HuH7 cells were not transfected as they stably express the fluorescently labeled actin. The HuH7 wild type cells were transfected with mRNA encoding for d2eGFP without using the perfusion system as the observation of the early fluorescence signal was not of importance to analyze the noise of the single-cell time courses. One six-channel array was observed per experiment with different combinations of two fabrication methods (three channels per method) to allow direct comparison with the same cell passage. All image stacks were background corrected.

Two different methods were used for the readout of the single-cell fluorescence time courses. Firstly, the mean fluorescence of each cell was calculated based on the cell contour at each time point. The fluorescence images were used for the determination of the cell contour as the contrast between cell and background is much better for the fluorescence images than for the phase contrast images. The main drawback of the method is that the contour detection is computationally intensive. However, the contour analysis can be used to determine the area of each single cell at each time point. The cell confinement was determined by calculating the area coefficient as the ratio of the cell area relative to the mean adhesion spot area. The mean adhesion spot area was determined from phase contrast images for each micropattern and kept constant for all cells measured on the respective micropattern. An area coefficient of one corresponds to the situation, where the cell area is equal to the adhesion spot area. Cell areas smaller than the adhesion spot area result in coefficients smaller than one and accordingly larger cells have a coefficient larger than one. Secondly, the mean fluorescence was calculated based on the pattern geometry using the MWA plugin as described in subsection 2.3.4..

3.2.1. Actin-eGFP expressing cells

In Figure 3.3. A the geometries of the three microarray fabrication methods are sketched. The adhesion spots of the 2-dimensional patterns for µPIPP and the gold pattern are squares with an edge length of 30 µm, which corresponds to an area of 900 µm^2. The cylindrical 3-dimensional µ-wells have a height of 30 µm and a diameter of 35 µm, which corresponds to an adhesion area of 962 µm^2. The single-cell time courses of the actin-eGFP signal are fluctuating around a mean value for most of the single-cell traces with a cell-to-cell variability regarding the mean fluorescence level (Figure 3.3.). The mean fluorescence time courses (Figure 3.3. B, green and orange traces) have an almost constant level for cells on the 2-dimensional pattern, as expected. The mean fluorescence time course of cells in the µ-wells (Figure 3.3. B, blue trace) showed a decreasing trend. This could be a hint that the cells are stressed by this confinement strategy in some way. The mean

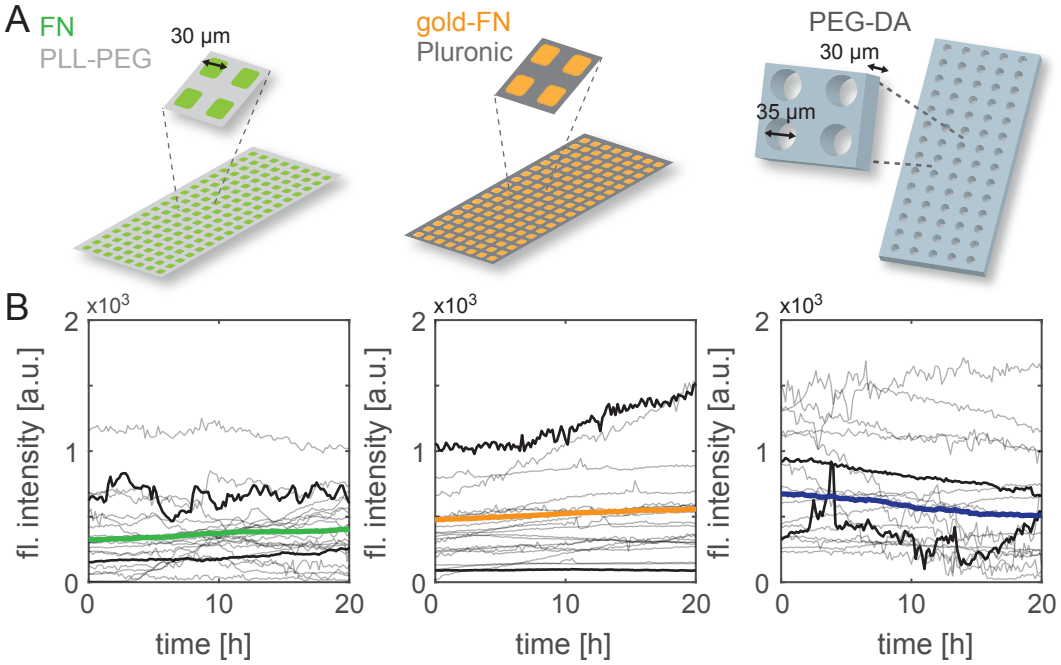

Figure 3.3.: Single-cell actin-eGFP time courses on the three different microarrays: µPIPP, gold pattern, or MIMIC µ-wells. (A) Sketch for each microstructure with respective geometries [M1]. (B) Representative single-cell time courses (gray) with the respective mean time course (colored) for each data set. The black traces show single-cell time courses with low or high noise.

fluorescence time course are in a similar range for all three microarrrays, which is a good indication that the expression is not affected by the choice of microarray. Representative traces of cells with a low or high noise level are highlighted. The traces with almost no noise indicate that the cell was well confined over the whole observation period, whereas noisy time courses most likely arise from cell area changes due to outgrow or shrinkage of the cell. Interestingly, the time courses of cells in the µ-wells are noisy although it was expected that the cells are perfectly confined inside the wells. The visual observation of

the time-lapse movies of cells in µ-wells showed that cells not only adhere on the bottom of the wells but also on the surface between wells or even the well's walls. Apparently the crosslinked PEG-DA structures are not cell repellent anymore. Furthermore, the PEG-DA has a high autofluorescence compared to PLL-g-PEG or Pluronic monolayers which led to problems for the background correction [96]. It often happened that pixels of a weak cell signal were identified as background due to the high fluorescence signals of PEG-DA. These two reasons lead to the noisy fluorescence time courses for µ-wells.

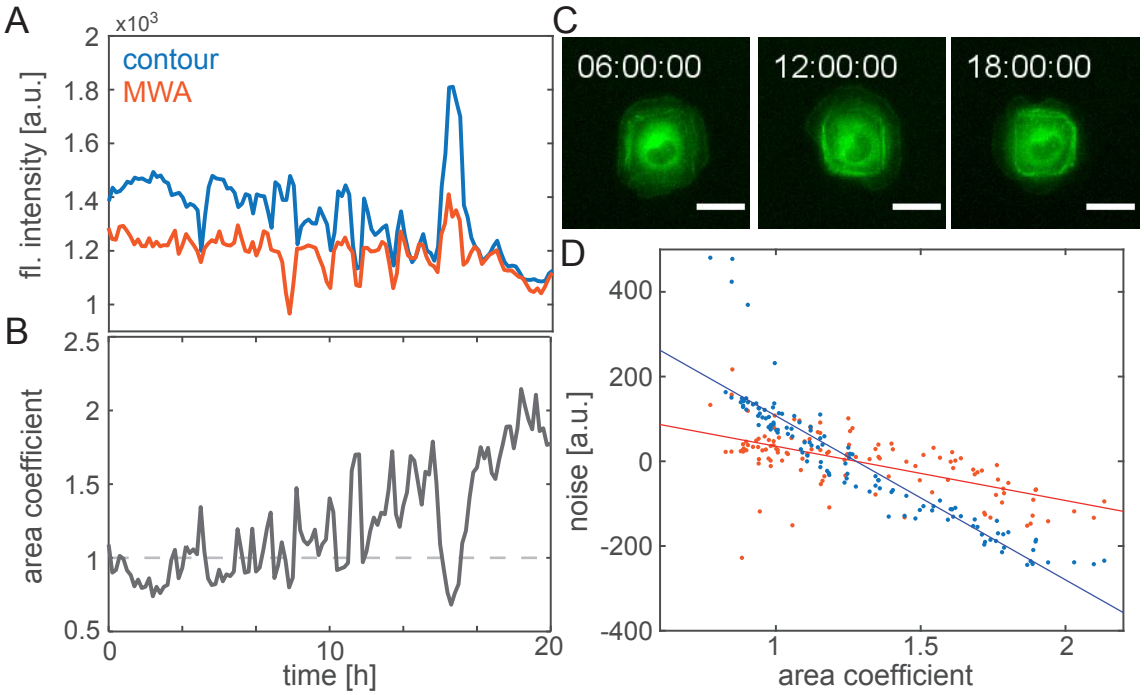

Figure 3.4.: Single-cell actin-eGFP expressing time course on a µPIPP adhesion spot. (A) Fluorescence intensity time course based on contour analysis (blue) or MWA (orange). (B) Time dependent changes of the area coefficient for the same cell. (C) Snap shots of the eGFP fluorescence images at three time points. The scale bars correspond to 25 µm. (D) Area coefficient vs. noise of the single-cell fluorescence time courses with respective linear regressions.

A representative single-cell fluorescence time course obtained from contour analysis (blue) or MWA (orange) on a FN square (µPIPP) is shown in Figure 3.4. A. The MWA time course appears to be less noisy than the contour-based time course. The respective area coefficient of the same cell is shown in Figure 3.4. B. For this particular cell the area coefficient tends to increase over time, which could be an indication that the cell confinement is time dependent. The fluorescence images of a representative cell shown in Figure 3.4. C at three different time points illustrate that the cell stays on the adhesion spot over time. However, the cell area changes over time as the cell outgrows of the adhesion spot. The representative single-cell data presented in Figure 3.4. illustrates what was observed for all cells independently on the microarray fabrication method. Fluctuations in intensity occur simultaneously with the change of the area coefficient. The noise of the fluorescence

time course can be traced back to the cell confinement of the cell to the micropattern. The area coefficient is negatively correlated to the noise of the fluorescence intensity time courses, which is shown by the linear regression for both time courses (Figure 3.4. D). The noise was determined by subtracting the mean fluorescence intensity, which is assumed as the biological expression level. In general, it was observed that the fluorescence time courses obtained by contour analysis had higher noise than the MWA time courses. For this reason, beside the computationally more intensive analysis of the contour analysis, the subsequent fluorescence intensities were obtained by the MWA approach.

It was shown that the noise is correlated with the cell area. For further analysis, the noise was autocorrelated. If the noise was random, the autocorrelation would be zero for each time delay larger than zero. In Figure 3.5. the autocorrelation function (ACF) of the noise time course of the fluorescence intensity time course obtained by MWA is shown. The ACF of the noise time course reveals that the noise is not random. It took more than two hours for the particular cell until the noise was uncorrelated. This is an indication that the temporal noise has a biological background.

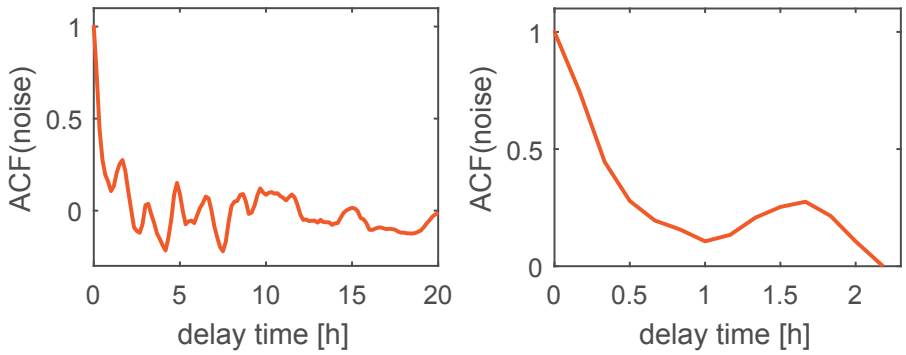

Figure 3.5.: Representative ACF of noise time course obtained from single-cell fluorescence intensity time course.

3.2.2. GFP expressing cells after mRNA transfection

The fluorescence time courses of actin-eGFP expressing cells showed that the noise can be attributed to changes of the cell area. In the subsequent studies of this thesis, the fluorescence intensity time courses of d2GFP expressing cells after mRNA transfection are investigated. For this reasons HuH7 wild type cells were transfected with L2000 lipoplexes containing mRNA as described in Appendix A. Representative time courses of d2eGFP expression after mRNA expression are shown in Figure 3.6.. The cells on all micropattern show a similar behavior with a high cell-to-cell variability regarding the fluorescence intensity, which corresponds to a high variation in protein level per cell.

Before the investigation of the time courses, the transfection efficiency was analyzed depending on the microarray fabrication method as the statistics for single-cell translation assays depend, among other things, on the number of d2eGFP positive cells. Therefore, the fraction of d2eGFP positive cells 15 h after transfection was determined for each of

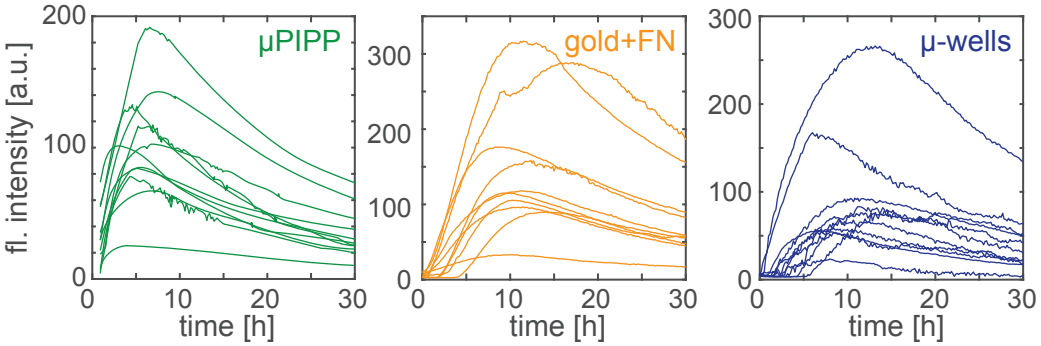

Figure 3.6.: Single-cell time courses of d2eGFP expression after mRNA transfection for cells on µPIPP (green), gold coated with FN (orange), or µ-well (blue) microarrays.

the three microarray methods as shown in Figure 3.7.. The fraction of d2eGFP positive cells to the total number of cells is defined as transfection efficiency (TE). The TE for cells seeded on the 2-dimensonal micropattern was on the order of 70-80% including the day-to-day variance between replicates with slightly higher TE for the gold micropattern. The TE for cells seeded in µ-wells dropped to 20% and it was observed that especially cells adhered to the bottom of µ-wells are rarely transfected. A possible explanation is that the L2000 lipolexes with their cationic surface charge stuck to the PEG-DA resulting in a lower TE for cells inside the wells. Due to the low statistics for cells in µ-wells only the data obtained from the 2-dimensional micropattern are further analyzed.

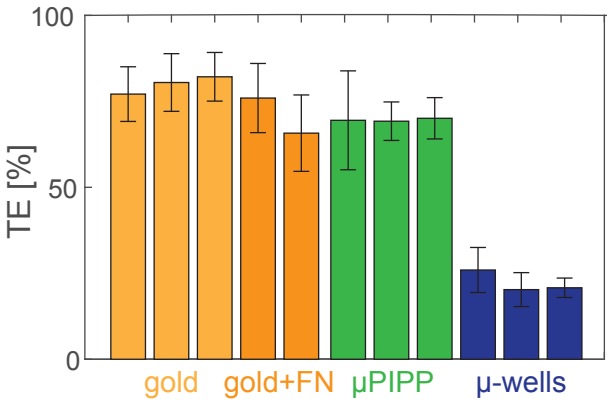

Figure 3.7.: The transfection efficiency TE as fraction of d2eGFP positive cells on different microarrays. The bar color indicates the microarray fabrication method and multiple bars of the same color correspond to different replicates. The error bars correspond to the standard deviation.

Single-cell protein expression time courses after mRNA transfection are analyzed to verify if the noise time courses are dependent on the micropattern and if they are autocorrelated. The noise time courses of each cell was calculated by fitting a translation model based on ODEs that was previously used to the experimental data [37]. In order to determine the noise, the residuals of the fit to the experiment time courses were calculated. The fitting approach and the corresponding model are explained in detail in the subsequent

chapters as the fit was used for noise calculation here. The single-cell time courses of cells in µ-wells were not further analyzed as the statistics were small due to the low TE and the background fluorescence of PEG-DA leading to increased noise during image processing. For the 2-dimensional patterns it was observed that the noise time courses of cells on µPIPP pattern is higher than the noise time courses on gold pattern (Figure 3.8. A). This is most likely due to larger cell areas on µPIPP pattern compared to cells on the gold pattern. The mean time courses were almost zero for times longer than 5 h which indicates that there is no technical noise source like illumination deviations. The deviations of the mean fluorescence until 5 h is model dependent and will be discussed in chapter 4. The autocorrelation of the noise time courses after mRNA transfection (Figure 3.8. A) shows a similar behavior as the representative autocorrelation of actin-eGFP noise time courses, shown in Figure 3.5.. Each of the single-cell time courses shows a certain degree of correlation but with a high variability.

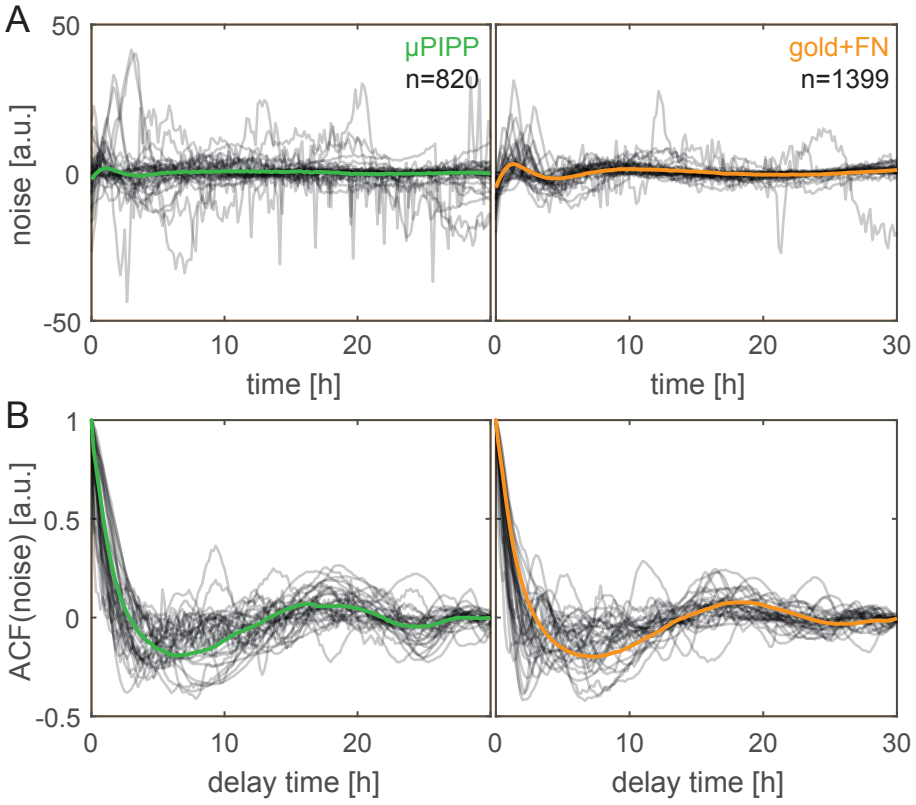

Figure 3.8.: Noise analysis of single-cell fluorescence time courses after mRNA transfection. (A) Representative subset of single-cell time courses of cells transfected on µPIPP pattern (left) or gold coated with FN (right). (B) ACF of the noise time courses of cells transfected on µPIPP pattern (left) or gold coated with FN (right). The gray time courses correspond to single-cells with the number of cells n and the respective mean time course (colored traces).

3.3. Discussion

It was possible to observe single-cell fluorescence time courses using each of the three described techniques. However, the fabrication methods have differences in how easy and reproducible they are to fabricate, in their ability to confine cells for longterm observations, and the fluorescence signal readout. The 2-dimensional micropattern appeared more suitable than the µ-wells because the background fluorescence did not interfere with the background correction. In addition to the image processing problems for µ-wells, the TE dropped for µ-wells compared to the 2-dimensional micropattern (Figure 3.7.). The µPIPP and the gold pattern lead to satisfactory results, whereby the gold pattern performed slightly better (Figure 3.8. A). The noise in single-cell fluorescence time courses proved to be correlated to the cell area (Figure 3.4. D) and thus to the cell's confinement on the micropattern. The autocorrelation of the noise time courses revealed that the noise is not uncorrelated in time. However, the mean autocorrelation cannot be used to determine a typical correlation length as the bulk dynamic arises due to superposition of the single-cell dynamics that is subject to a large cell-to-cell variability. This phenomenon was previously observed for single-cell dynamics [65]. The single-cell noise correlation might arise from cell motility as the time scale in the range of 10 min up to an hour corresponds to a similar time scale of lamellipodium formation (protrusions of cells related to migration) [97]. In order to verify if the noise correlation is dependent on cell internal mechanisms, further investigations are necessary, which is beyond the scope of this thesis.

The use of the 3-dimensional MIMIC pattern with PEG-DA proved to be unsuitable for the readout of fluorescence time courses. Beside the low TE and the high background fluorescence, the cells did not preferentially adhere on the bottom of the wells. In order to functionalize the bottom of the µ-wells, which could increase cell alignment, a combination of protein printing (described in [73]) and the MIMIC apporach was tested. Therefore, the stamps were pre-incubated with a FN solution before being placed on the coverslips. The remaining MIMIC protocol was kept the same. The idea is to transfer FN onto the coverslip and thus functionalize the bottom of the µ-wells to increase adhesion. The transfer of FN was checked by immunofluorescence staining. A primary antibody against FN was incubated to enable binding to FN, followed by an incubation of a second fluorescently labeled antibody specific to the FN antibody. Figure 3.9. shows a fluorescence intensity profile of µ-wells with printed FN (purple) and µ-wells without FN (orange, the second orange trace corresponds to a control where only the second antibody was incubated). It is shown that the µ-wells with printed FN have a much higher fluorescence signal on the bottoms of the wells compared to the PEG-DA structure. The opposite is the case for µ-wells without FN. Nevertheless, the fluorescence background remains a problem besides the low transfection efficiency compared to the 2-dimensional micropattern techniques.

The gold pattern proved to be a suitable method for single-cell microarray fabrication. The passivation with Pluronic-F127 showed superior passivation compared to PLL-g-PEG as reported in literature [98] resulting in a more stable cell confinement (Figure 3.8. A). Additionally, the approach is highly reproducible. The fabrication needs to be performed in a clean room which is less practicable compared to the other methods presented and the fluorescence signal is attenuated as the illumination occurs through the metal layers. It

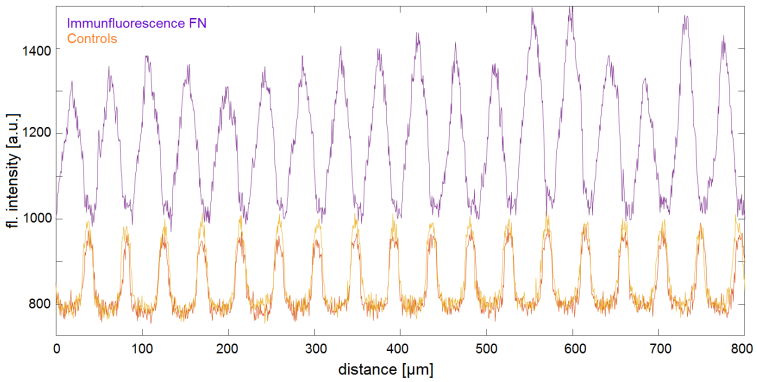

Figure 3.9.: µ-wells with FN functionalized well bottoms. Intensity profile of 18 µ-well structures. The presence of FN was verified by secondary immunofluorescence staining against FN (purple). The control intensity profiles of µ-wells without FN or only incubated with labeled antibody showed higher fluorescence intensity for the PEG-DA walls than on the bottom of the wells (orange).

was further observed that a coating of the gold squares was necessary in order to obtain a similar adhesion behavior on the gold pattern compared to the other two methods. However, the method is particularly interesting for sub-structuring adhesion spots like illustrated in Figure 3.10. showing a HuH7 cell on a gold square that is sub-structured by small gold islands over a period of 3:20 h. The phase contrast images illustrate the good cell confinement on the gold structures. Eventually, the protrusion dynamics can be investigated in more detail using different substructure geometries. This could have an influence on the protrusion formation and increase cell confinement beside the better passivation of Pluronic.

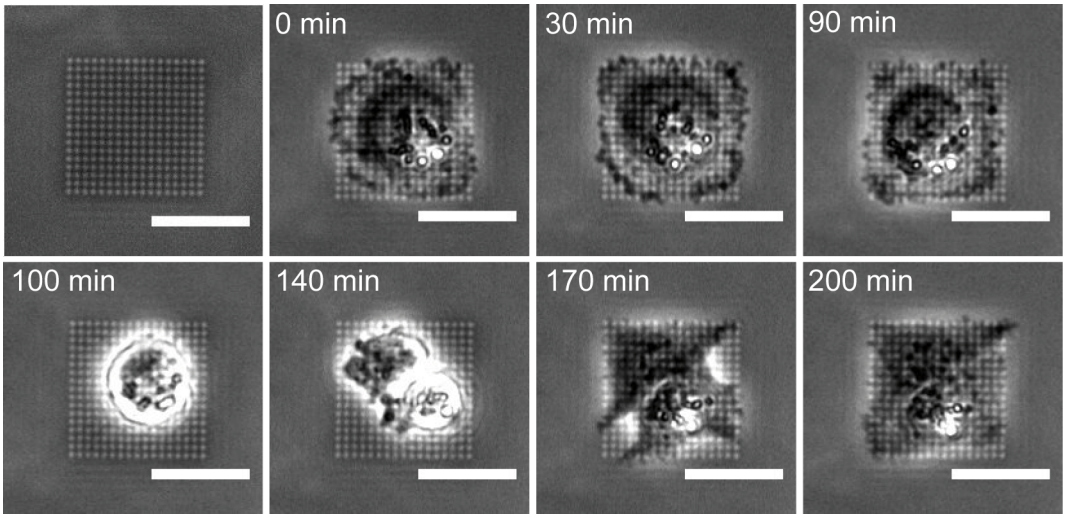

Figure 3.10.: Dividing HuH7 cell on sub-structured gold adhesion square with an edge length of 30 µm. The scale bars correspond to 25 µm.

The µPIPP arrays are relatively simple to produce and high statistics for single-cell flu-

orescence time courses are achieved. A drawback of the pattern technique is that the pattern quality especially for passivation showed day-to-day variability. The noise for the µPIPP time courses is only slightly higher than for the gold time courses. The autocorrelation for both sets of time courses were very similar, which is a further indication that the correlation occurs due to a biological process that is independent of the micropattern. For these reasons, the µPIPP approach was used for all further studies discussed in this thesis.

4. Multi-experiment nonlinear mixed effect modeling of translation kinetics

In systems biology, mechanistic pathway models based on ODEs (see section 2.2.) are a standard tool to quantify heterogeneity of cellular reaction networks to obtain a better understanding of biological processes [4]. However, the estimated parameters depend on the selected model and are limited if the estimation is only related to one kind of experiments, which is the case for most single-cell assays [83]. The integration of data collected under varied conditions is expected to improve parameter estimation and thus facilitates the gain of knowledge. In order to integrate several single-cell experiments for more robust parameter estimation, new approaches need to be developed. Here, perturbation experiments using the already described single-cell translation assay (see section 2.3.) were performed and integrated by a novel approach based on multi-experiment nonlinear mixed effect modeling. The idea of the perturbation experiments is to systematically vary one parameter of the investigated biological process and keeping the others constant. For the translation kinetics after transfection that is achieved by changing the sequence of the transfected mRNA. As a proof of concept, two mRNA constructs with the same structure but different ORF sequences were transfected and the translation kinetics of both constructs were assayed under the same conditions. The sequence of the two investigated ORFs encode for eGFP or a destabilized version of eGFP called d2eGFP, which has a lower half-life than eGFP [99].

In order to proof if perturbation experiments are beneficial, two commonly used fitting approaches were compared. The (I) standard two-stage (STS) and the (II) non-linear mixed effect (NLME) approach were used to estimate the translation model parameters of the obtained fluorescence trajectories:

(I) STS With the STS approach, as described in section 2.4., the single-cell parameters and the respective population distributions are calculated one by one [100, 101]. An ODE based model is fitted independently to each single-cell trajectory to estimate the parameter set for each cell. Afterwards, the parameter distribution for the population is reconstructed using the single-cell parameters. The approach is easy to implement and therefore often used [21, 33, 102]. The drawback of the method is that the cell-to-cell variability is overestimated because uncertainty of the estimated parameters and biological heterogeneity cannot be distinguished [103].

(II) NLME The single-cell parameters and the population distribution are estimated si-

multaneously for the NLME approach [104]. The single-cell parameters are considered as variables that are constrained by the population distribution of each variable. The approach is in general computationally more elaborated than the STS. The method was recently established for fitting single-cell data [100, 101, 105, 106]. NLME reduces parameter uncertainty [101, 103] and has less estimation bias [100] compared to STS.

In this study, it could be confirmed that the NLME improves parameter estimation compared to STS for the analysis of single-cell translation kinetics. The use of perturbation experiments by the integration of the two GFP variants was used for an extension of the NLME to a multi-experiment NLME that improved the parameter estimation even more. The multi-experiment NLME was further used to investigate four different translation models. This new method presented the evidence that ribosomal activity is rate limiting for translation. Moreover, the approach allows a robust parameter estimation across experimental replicates, which enables mechanistic insights into the nature of batch effects. The study was performed in cooperation with members of Jan Hasenauer's group (Helmholtz center, Munich), who developed the employed fitting algorithms. The chapter is based on the publication [P2] listed in Associated publications and manuscripts. The data and the affiliated algorithms are deposited at zenodo (doi:10.5281/zenodo.1228899).

4.1. Translation kinetics after transfection show large cell-to-cell variability

The fluorescence time courses were obtained by the single-cell translation assay, which was described in section 2.3.. HuH7 cells were transfected with mRNA-L2000 lipoplexes during the first hour of the experiment using two different mRNA constructs encoding for eGFP and d2eGFP respectively. The kinetics were observed over a duration of 30 h. Cooperation partner from ethris GmbH synthesized the two mRNA constructs, which were used in this study, with the ORF of eGFP and d2eGFP cloned in the same vector backbone pVAX1-A120 as described previously [107]. Three independent replicates of the single-cell translation assay were performed to examine the reproducibility of the assay and if the parameter estimation is robust regarding to day-to-day variances. A total number of ∼2600 cell trajectories for eGFP and ∼2900 cells for d2eGFP were measured. These numbers correspond to an average of roughly 350 trajectories per microarray channel as visualized in Figure 2.4.. The yield of fluorescence trajectories depend on several factors like the micropattern quality, the generation time of the used cell line, or the experiment duration.

A large cell-to-cell variability of the translation kinetics was observed for both constructs in all three replicates (Figure 4.1.). The fluorescence time courses of each cell originated from the translation of fluorescence proteins after transfection with mRNA using L2000 lipoplexes (Figure 4.1. A). The heterogeneity of the fluorescence time courses occur from a variability of all parameters describing the kinetics like the degradation rates or the translation rate. The visual inspection of the fluorescence time courses of eGFP and d2eGFP expressing cells already shows qualitatively different translation kinetics for both

mRNA constructs (Figure 4.1. B). The mean time courses (black) of eGFP and d2eGFP both peak around 10 h. The eGFP signal stayed almost constant afterwards, whereas the d2eGFP signal declined subsequently. The faster decrease of the fluorescence signal for d2eGFP expressing cells indicate a higher protein degradation rate of d2eGFP compared to eGFP as expected from the literature [99, 108].

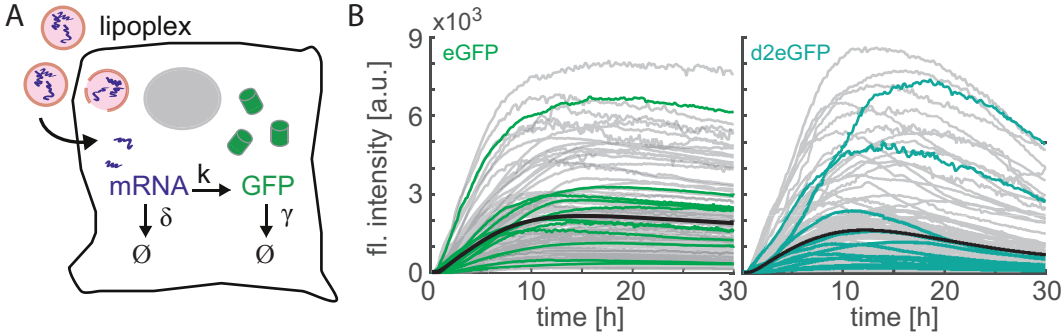

Figure 4.1.: Single-cell perturbation experiments using two different GFP variants. (A) Sketch of mRNA lipofection of a cell. The mRNA-lipoplexes are taken up by the cell. After mRNA release, the mRNA is translated into GFP with rate k. mRNA and GFP undergo degradation with rate δ and γ respectively. The protein degradation rate γ is systematically varied using two different GFP variants. (B) The single-cell time courses reveal qualitatively different translation kinetics of eGFP (left plot) compared to d2eGFP (right plot). A representative subset of fluorescence time courses (gray traces) are shown with the respective mean trace (black) for each data set. The green time courses highlight individual time courses.

Previous studies already revealed the high cell-to-cell variability of protein expression after mRNA transfection [21, 33] suggesting stochastic mRNA delivery events [20] and limited enzyme abundances e.g. for ribosomes [109]. The translation model used in the previous transfection studies is an two-stage model, as visualized in Figure 4.1. A, and describes the temporal relation between $mRNA$ and GFP concentration after mRNA transfection:

$$\begin{aligned} \frac{d[mRNA]}{dt} &= -\delta[mRNA],\ [mRNA](t_0) = m_0 \\ \frac{d[GFP]}{dt} &= k[mRNA] - \gamma[GFP],\ [GFP](t_0) = 0 \end{aligned} \tag{4.1}$$

with k being the translation rate per mRNA molecule, m_0 as amount of transfected mRNAs released into the cytosol at time point t_0, δ is the mRNA degradation rate, and γ is the protein degradation rate. The measured fluorescence intensity y is defined as the logarithm (due to computational reasons) of the GFP amount multiplied by scaling factor *scale* plus the background fluorescence *offset* (analog to Equation 2.3):

$$y = log(scale[GFP] + offset) \tag{4.2}$$

The analytic solution of the above ODE system (Equation 4.1) for the translation model is:

$$[GFP](t) = \frac{m_0 k}{\delta - \gamma}(e^{-\gamma(t-t_0)} - e^{-\delta(t-t_0)}) \tag{4.3}$$

The analytical solution (Equation 4.3) reveals the structural non-identifiability of both degradation rates, the number of transfected mRNA molecules, and the translation rate as previously described [21]. Therefore, the model is transformed to reduce the number of free parameters, resulting in a combination of parameters as one fitting parameter i.e. km_0scale. The transformation was done by normalizing the mRNA concentration by m_0 and the GFP concentration by the scaling factor. Details are given in the methods section and the supplementary information of [P2]. The degradation rates can only be distinguished for this model if assumption are made e.g. that the protein degradation rate is smaller than the mRNA degradation rate. This assumption cannot always be made and depend on the properties of the IVT mRNA and the stability of the expressed reporter protein.

4.2. Comparison of fitting approaches

4.2.1. Standard two-stage approach

The eGFP data set was fitted using the STS like illustrated in Figure 4.2. A to estimate the parameters of the translation model given by the ODEs in Equation 4.1. In stage 1 the parameter set of an individual cell was determined using a MLE method followed by the estimation of the parameter distribution of the cell population (stage 2) like described in literature [100, 101]. For computational reasons the single-cell time-courses are logarithmized (Figure 4.2. B) and the model fits describe the data reasonably well as shown in Figure 4.2. C.

The degradation rates for mRNA δ and protein γ cover two orders of magnitude and indicate multiple modes (Figure 4.2. D). One could conclude that the multiple modes are caused by subpopulations. Yet, the modes are caused by the parameter symmetry of δ and γ (as can be seen in Equation 4.3) and revealed by an uncertainty analysis of the single-cell parameters (see methods section of [P2]). The values for the degradation rates can be interchanged without changing the translation kinetics. In Figure 4.2. D the estimated population distributions are shown if the degradation rates are swapped. The symmetry of the degradation rates lead to a global structural non-identifiability [110]. But, a structural identifiability analysis showed that the parameters could be resolved by measuring the mRNA and the protein concentration of each cell or by measuring two protein levels by transfecting two mRNA constructs coding for different proteins. Both approaches are conceivable but non-trivial [111, 112]. Possible implementation of both methods for the single-cell translation assay are discussed in chapter 7. In general, an unique identification of the parameters revealed from single-cell experiments is often not

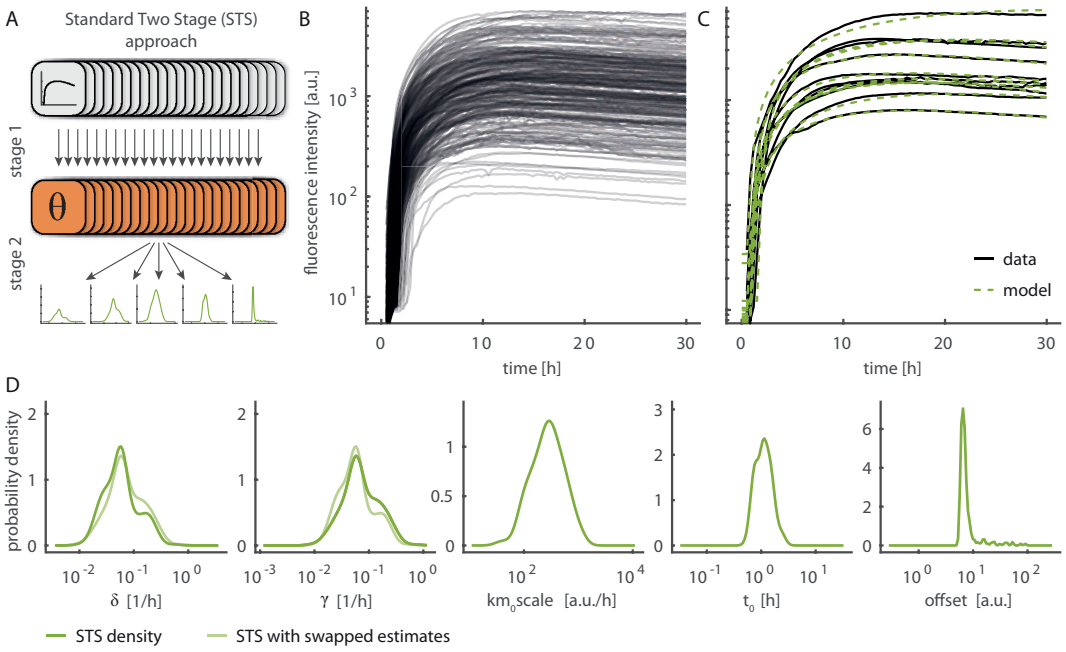

Figure 4.2.: Results of the parameter estimation for the STS approach. The STS approach is based on estimating the parameter set for every trace (stage 1) and calculates the population distiribution for each parameter (stage 2) after fitting each trace (A). Single-cell trajectories of eGFP expressing cells (B) and ten trajectories with their corresponding fits (C). The parameter distributions are computed using a kernel density estimation (D). Adapted from [P2], licensed under creative commons.

possible by established experimental methods. For population experiments it is reported that perturbation experiments can solve structural non-identifiabilities [113]. The perturbation experiments are used as additional observations, which are integrated in the model [114]. This is possible under the assumption that the model parameters are conserved quantities and do not change between different experiments. That assumptions makes it difficult to consider perturbation experiments for STS, as the parameters for all experiments are assumed to be independent, which exclude conserved quantities.

4.2.2. Non-linear mixed effect approach

As STS is not suitable to resolve the structural non-identifiabilitiy of the degradation rates, a NLME was applied to consider perturbation experiments. The NLME allows for conserved quantities in perturbation experiments. The single-cell parameters φ_i of the NLME are assumed to consist of a fixed effect β and a normally distributed effect $b_i \sim \mathcal{N}(0, D)$: $\varphi_i = exp(\beta + b_i)$ as visualized in Figure 4.3. A. The population parameters β and D are the mean and the covariance of the logarithmic single-cell parameters. It is assumed that the population parameters are the same for different experimental conditions and even for individual cells. As a result, the single-cell time courses of cells transfected with different mRNA constructs can be integrated at the population level. The single-

cell parameters and the population parameters are determined in parallel (for detailed information see the methods section for the NLME approach in [P2]).

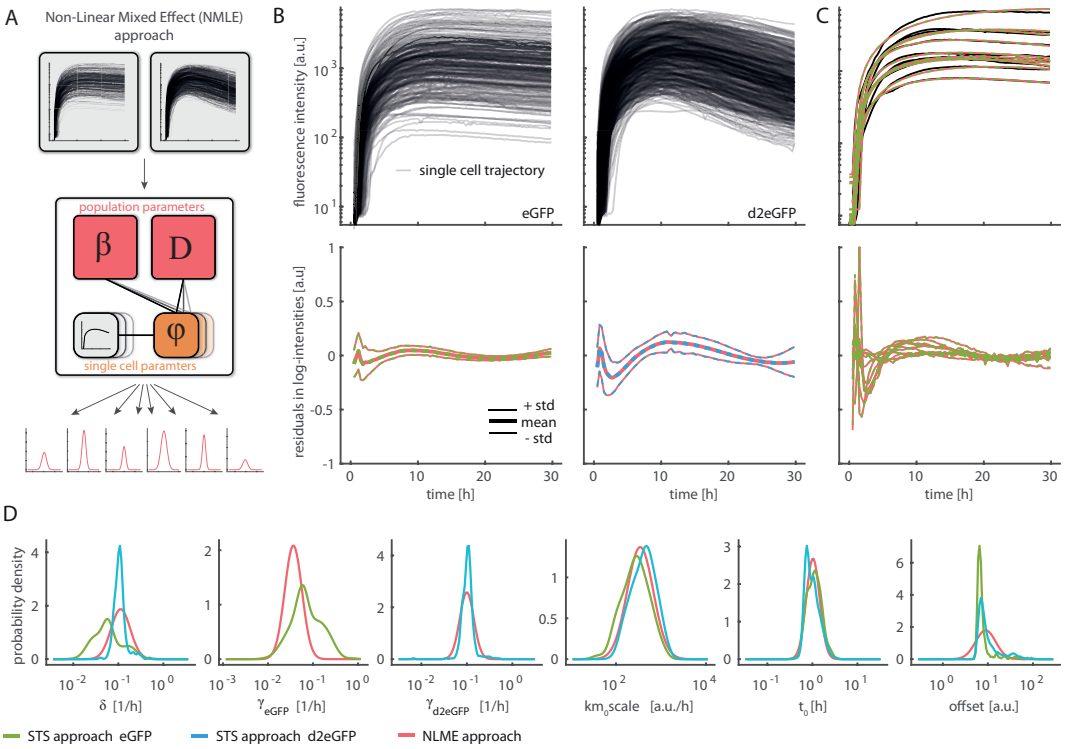

Figure 4.3.: NLME and STS compared regarding the estimation of the parameter distributions. (A) Sketch of the NLME. (B) Representative time courses of eGFP and d2eGFP (top) and the mean residuals of both approaches with respective standard deviation (std) (bottom). (C) Ten representative single-cell fits (top) for eGFP expressing cells with respective residuals (bottom). (D) Estimated parameter distributions of both approaches. Coloring indicate the applied approach and in combination with the data set (STS of eGFP in green, STS of d2eGFP in blue, and NLME in red). Adapted from [P2], licensed under creative commons.

In order to examine if the degradation rates are identifiable using perturbation experiments, the data sets of eGFP and d2eGFP expressing cells (Figure 4.3. B) were combined and the NLME was extended to a multi-experiment approach (see respective method section of [P2]). For the NLME based fitting two distinct distributions for the protein degradation rates were assumed for eGFP γ_{eGFP} and d2eGFP γ_{d2eGFP}. All other parameters are assumed to be the same for the perturbation experiments. For STS it is also assumed that the protein degradation rates differ and the other parameters are the same for both data sets. However, it cannot be enforced that the other parameter distributions, like the mRNA degradation rate, are the same for the two perturbation experiments using STS.

Both approaches yield very similar fits of the single-cell time courses as can be seen for the almost identical population statistics of the residuals in Figure 4.3. B and for representative single-cell fits and the respective residuals in Figure 4.3. C. For multi-experiment

NLME, the estimation of the single-cell parameters φ_i and the estimation of the population parameters β and D agreed with each other and describe the data reasonably well, which suggest that the assumptions for the NLME were appropriate (Figure 4.4.). The population distribution for both fitting approaches results in comparable distributions (Figure 4.3. D) even for γ_{d2eGFP}. The reason why multiple modes for γ_{d2eGFP} are not visible for STS compared to γ_{eGFP} can be explained due to the fact that the mRNA and d2eGFP degradation rates are in a similar range. For this reason, the symmetry effect of the mRNA and d2eGFP degeneration rates are neglectable. The STS approach results in considerable different distributions for the mRNA degradation rate δ depending on the data set, while the multi-experiment NLME yields one narrower distribution compared to STS (Figure 4.3. D). The smaller distribution for multi-experiment NLME can be traced back to the symmetry breaking of the degradation parameters by the consideration of an additional perturbation experiment (see supplementary information of [P2]). The single-experiment NLME approach was not capable to break the symmetry of δ and γ and is for that reason not further discussed (Figure S2. of [P2]). These results demonstrate that multi-experiment NLME is suitable to consider perturbation experiments, which leads to improved single-cell parameter estimation compared to STS.

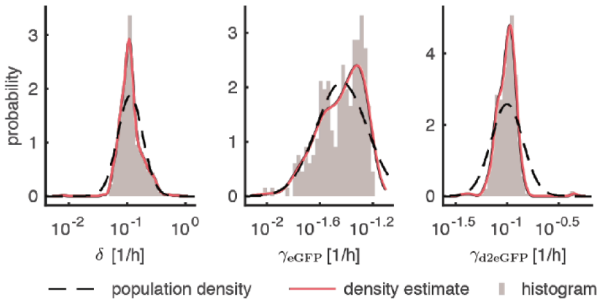

Figure 4.4.: Comparison of distributions of estimated single-cell parameters φ_i and estimated population parameter distribution (β, D). Multi-experiment NLME single-cell parameter histograms and respective kernel density estimate (red line) with the population constraint (black dashed line) shown for the degradation rates. Adapted from [P2], licensed under creative commons.

4.3. Translation model selection

Yet, the suitability of the translation model needs to be questioned due to the systematic variations between the fits and the single-cell time courses for STS and NLME, which can be seen by the mean residuals in Figure 4.3. B (bottom). The residuals for both data sets reveal a clear temporal trend, indicating that the model is inappropriate to describe the experimental data. For this reason, additional translation models were analyzed in which some of the kinetic rates of the translation model are no longer described as first order reactions. The assumption of first order kinetics is only true if the abundance of the affected molecules is not rate limiting.

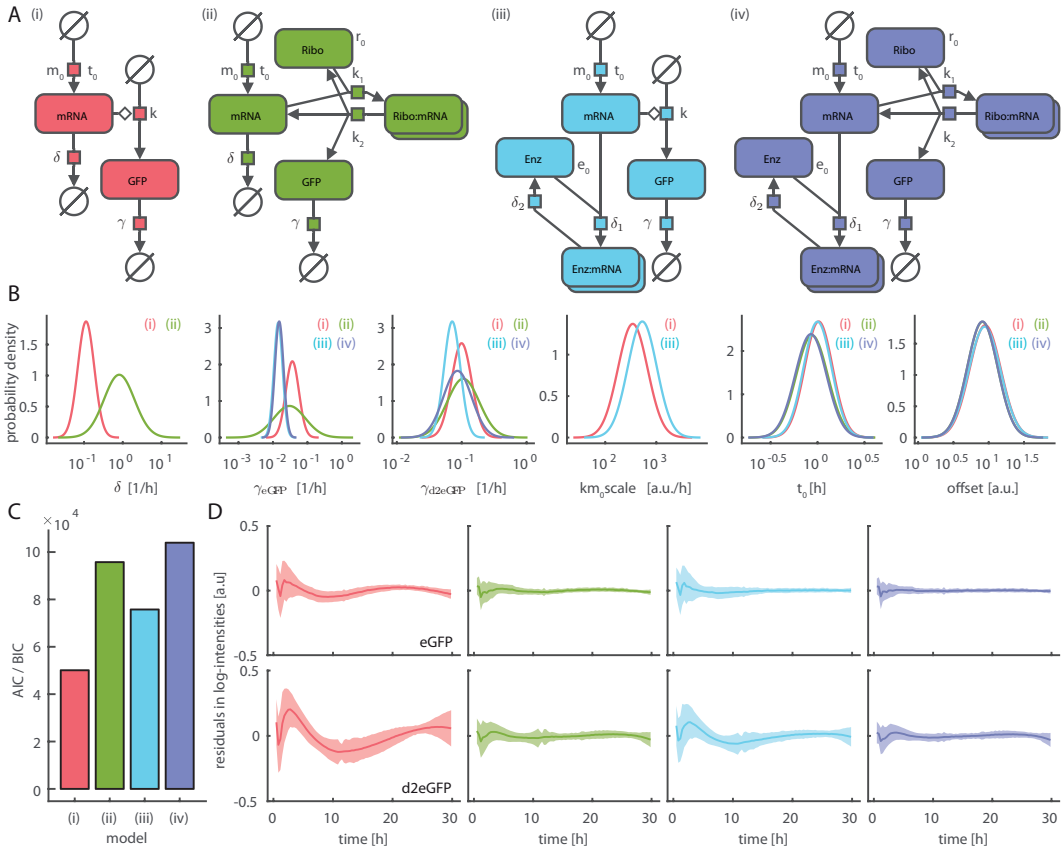

Figure 4.5.: Four translation models are compared regarding the parameter estimation. The coloring corresponds to the model. (A) The model candidates are plotted using systems biology graphical notation (SBGN) [115]. (B) The parameter distribution across all four models are compared to each other using multi-experiment NLME. The model number is added to each parameter distribution if the model contains this parameter. (C) The four models are compared regarding their AIC/BIC value. (D) The residuals (mean ±std) for the eGFP and d2eGFP data sets are plotted for each model. Adapted from [P2], licensed under creative commons.

In total, four translation models were compared and their respective biochemical networks are illustrated in Figure 4.5. A. The least complex and already discussed translation model (i) is described by the ODEs (Equation 4.1). The ODEs of the three other models are given in Appendix B. The ribosome model (ii) describes the translation step by taking ribosome binding to the mRNA into account for the expression of protein. The enzyme degradation model (iii) describes the mRNA depletion depending on enzymatic activity. Model (iv) is a combination of model (ii) and (iii). The extensions of the models were chosen due to experimental evidence obtained in other studies [116]. An enzymatic degradation of the protein was also considered. However, it has been reported that fluorescent reporter proteins, like eGFP, are normally very stable and resistant to enzymatic degradation and can be described with a first order degradation kinetic [99]. For d2eGFP this is not necessarily the case because of the addition of a C-terminal PEST sequence, which

is naturally found in short-lived proteins and is known to be important for fast protein turn-over [108]. Li et al. [117] showed that eGFP is stable even in presence of the translation inhibitor cycloheximide (CHX) but that d2eGFP is very unstable due to the PEST sequence. To prove, that in our experimental set up the protein degradation of both reporter proteins can be described by first order kinetics, a translation block experiment for cells expressing d2eGFP was performed as described in the appendix and visualized in Figure A.4. [118, 119]. The fluorescence protein degradation was investigated after adding CHX 7 h post transfection. A simple exponential decay provided a good fit to all observed fluorescence intensity time courses, implying the sufficiency of first order kinetics in the considered concentration range (see Figure S1 [P2]). The described extensions for translation and mRNA degradation have not been studied for transfection dynamics and it is unclear if the other models lead to significant improvement of the fits.

The ribosome binding is modeled by an initial ribosome abundance R_0, which bind to $mRNA$ with rate k_1. GFP is translated by rate k_2 from the complex $Ribo:mRNA$ and produces free $mRNA$. The enzymatic mRNA degradation is described by introducing degradation enzymes with an initial abundance E_0. The degradation is replaced by a binding reaction of the enzyme Enz to $mRNA$ with rate δ_1 forming a complex $Enz:mRNA$. The mRNA degradation of the bound molecule takes place with rate δ_2 releasing the enzyme. For the enzyme degradation-ribosome model (iv), the translation and mRNA degradation are replaced by both of the described binding reactions. The parameters of all four models were estimated using the described multi-experiment NLME approach. The parameter distribution of the translation models reveal differences, especially for degradation rates δ, γ_{eGFP}, and γ_{d2eGFP} (Figure 4.5. B). These differences stress the need of appropriate mechanistic models when estimating single-cell parameters.

In order to identify the most suitable model candidate, two model selection criteria were considered: the Akaike Information Criterion (AIC) [120] and Bayesian Information Criterion (BIC) [121]. Both criteria favor the models with ribosomal activity (model (ii) and (iv)) like illustrated in Figure 4.5. C. Model (iv) achieved the best AIC and BIC values, which is the most complex of the investigated models taking ribosome and enzyme binding reactions into account. Beside the model selection criteria the residuals of all models were computed as shown in Figure 4.5. D. The residuals' magnitude of the models with ribosomal translation (model (ii) and (iv)) are substanially smaller than for the other two models. This indicates that the approximation of the translation process by a first order kinetic is not suitable and that the concentration of free ribosomes is rate-limiting for protein expression. The enzymatic degradation reaction in contrast is not essential for describing the data, as the residuals' differences between model (ii) and (iv) are small. For these reasons and due to the smaller number of free parameters, the ribosome model (ii) was chosen for the subsequent analysis.

4.4. Ribosome model identifies and explains batch effects

The setup of the single-cell translation assay allowed to study both perturbation conditions within the same experimental batch by transfecting cells of different microarray channels

with one of the mRNA constructs encoding for eGFP or d2eGFP, respectively. It is not always possible for an experimental setup and often the different perturbation experiments have to be measured using different batches. The use of different batches, like cells of a different passage, makes it difficult to distinguish batch effects from perturbation effects. For many single-cell approaches the data of different batches need to be pooled for higher statistics [105], which complicates to differ between cell-to-cell variability and batch effects. For pooling multiple perturbation experiments, it is necessary that the parameter estimation is robust regarding batch effects.

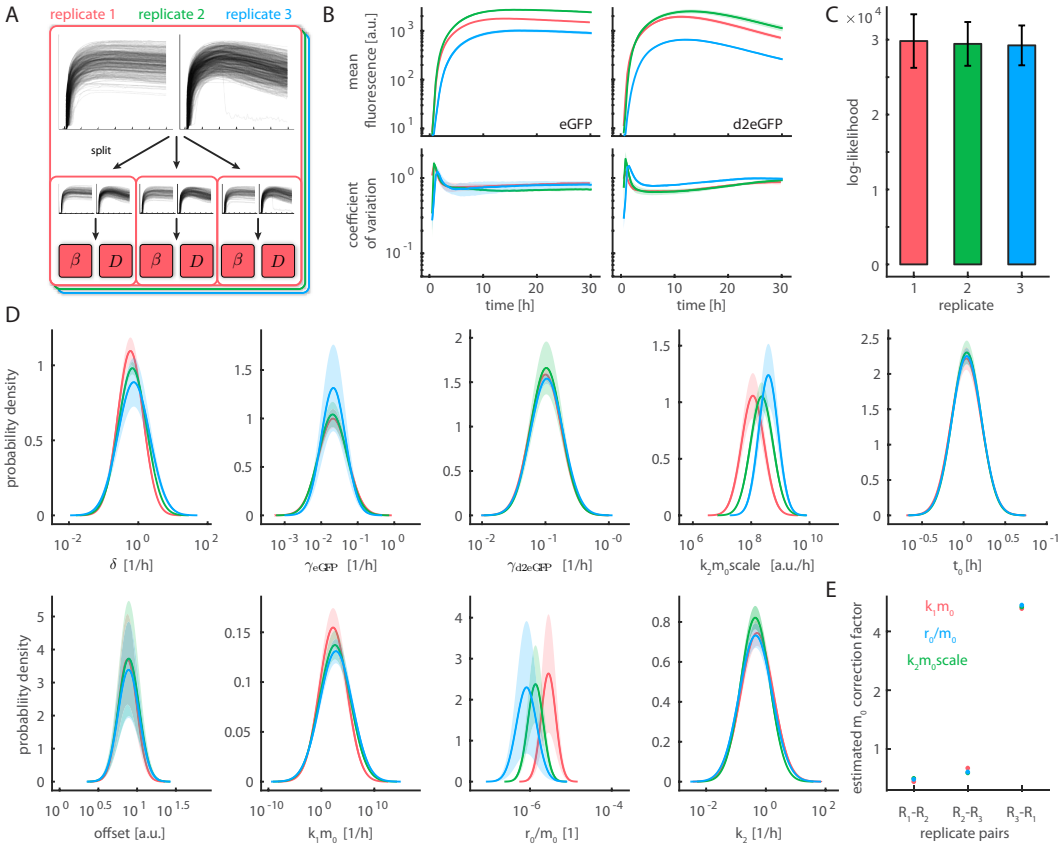

Figure 4.6.: Uncertainty analysis of replicate effects using the ribosome model. Coloring is according to the replicate. (A) Sketch of the three replicates and the corresponding threefold split of each replicate. (B) Variability within and across replicates given by the mean time courses and the coefficient of variation. The shaded areas indicate the sampling error. (C) Log-likelihood values within and across replicates. (D) Uncertainty analysis of parameter distributions within and across replicates. (E) Correction factor for m_0 regarding the estimated parameter distributions. All errors (shaded areas and bars) correspond to ±std. Adapted from [P2], licensed under creative commons.

In order to investigate if the parameter estimation of the multi-experiment NLME is robust and reproducible, three experimental replicates were analyzed. Furthermore, each replicate was divided into three subsets of the same size, leading to nine data sets as visualized in Figure 4.6. A. The mean single-cell time courses of each replicate show varia-

tions between the replicates, which are larger than the sampling error of the three subsets (Figure 4.6. B). To evaluate the influence of these batch effects on parameter estimation, each of the nine data subsets were analyzed individually. The log-likelihood values of the replicates were very similar, which indicates that the quality of the fitting was comparable (Figure 4.6. C).

The estimated parameter distributions were consistent within each replicate and across replicates in most instances (Figure 4.6. D). The distributions with observable differences across replicates were $k_2 m_0 scale$, $k_1 m_0$, and R_0/m_0. The mean of the population differed essentially across replicates for these parameters. Interestingly, all three parameter combinations share one parameter, the initial amount of mRNA molecules m_0. The parameter m_0 describes the average number of released mRNA molecules into the cytosol of a cell after transfection. Since the efficiency of mRNA delivery depend on a variety of things, like the cell culture conditions or lipoplex formation, it is likely that differences between replicates for the amount of delivered mRNA molecules are larger than the intrinsic distribution of mRNA abundance of one experiment. Figure 4.6. E shows that the correction factors for m_0 are consistent for the discussed parameters and all three replicates. It is possible, that the observed batch effects could be traced back on differences in the average number of released mRNAs. In this case, the multi-experiment NLME allows for mechanistic insight in the emergence and identification of batch effects.

4.5. Discussion

Scanning time-lapse measurements are capable to quantify and unravel cellular kinetics at the single-cell level. NLME proofed to be a powerful statistical approach to analyze fluorescence time courses, which is in agreement with previous studies [100, 101, 105, 106]. The integration of perturbation experiments using NLME for single-cell measurements was applied, to the best of my knowledge, for the first time. In this study, the expression kinetics of a fluorescent reporter after mRNA transfection was studied with single-cell resolution. As perturbation experiments, mRNA constructs encoding for GFP variants with unequal degradation rates were employed. Fluorescence intensity time courses for both GFP variants where measured in parallel for hundreds of cells per experiment. The inclusion of both data sets into a multi-experiment NLME yield to a more robust parameter estimation than for the NLME of each data set by itself and even better compared to the STS. In particular, the identification of the degradation rates for mRNA and protein is now possible for the translation model (i) (Figure 4.3. D) without prior assumption of the relation between both rates. Interestingly, the integration of perturbation experiments on the population level led to an improvement of the single-cell parameter estimation (see Figure S2. of the supplementary information [P2]). Hence, the multi-experiment extension is an important improvement. Generally speaking, the iteration of experiments using the same readout under systematically varied conditions alters the entire parameter set and does not improve the estimation quality.

For the less complex translation model (Figure 4.5. A (i)) a pronounced structure of the residuals was identified (Figure 4.3 B, bottom). The temporal trend of the residuals dis-

closed the limitations of the model for suitable description of the translation process. The comparison of four model candidates showed that the ribosome binding reaction lead to a higher agreement of the respective models with the experimental data (Figure 4.5. D). Indicating that the abundance of free ribosomes is rate limiting leading to another source of heterogeneity due to stochastic variations of ribosome abundance in cells [122]. However, the residuals for early time points around the protein expression onset still shows higher magnitudes compared to later time points (Figure 4.5. D). The experimental single-cell time courses show smoother onset behavior than the fits of each model, which could be solved by including a GFP maturation step or another more complex step like multiple mRNA release events to distinguishable time points.

Some of the population distributions show a large variability, for example for $k_1 m_0$ as shown in Figure 4.6. D. This could result form combining variables, which both underlay a high variability. Another explanation could be that the model for the transfection or translation process is too simple. This lead to an accumulation of the variability along a process of multiple steps resulting in a high variability of a few parameters [55]. The results of these study depend on the chosen mechanistic model comparable to similar studies. More detailed models might improve the parameter estimation but the increasing number of parameters need to be examined carefully. It is to mention, that the determination of the mean protein half-lives of 22.8 h for eGFP and 6.6 h for d2eGFP are in good agreement with the literature [99]. Due to this agreement, it is concluded that the multi-experiment NLME lead to reliable results.

Furthermore, the parameter estimation was reproducible for independent experimental replicates (Figure 4.6.). However, batch effects could not be prevented between replicates. The batch effects are often not taken into account for single-cell measurements and are not handled properly for pooled data leading to an interaction of batch effects with the biological cell-to-cell variability. Using the multi-experiment NLME, it was possible to identify sources of batch effects, which most likely occur due to a high variation of the average number of released mRNAs for different replicates. The identification of sources for batch effects could be used to further improve the experimental set up and therefore data quality.

In conclusion, the novel approach provides a powerful tool for parameter estimation of single-cell data. The fact that the integration of perturbation experiments are beneficial for parameter estimation and that the estimation gets even more robust enables new opportunities for the analysis of translation kinetics depending on systematic sequence modification. This is of interest for gene therapeutic purposes, as it enables to quantify the expression dynamics under varied conditions [123, 124]. Hence, an improved understanding of the expression kinetics and how the expression can be controlled is important for successful treatment design [14]. The only drawback of the novel fitting approach is the very high computational time which is needed for parameter estimation. For the presented results the total computation time took in total over 15 years cpu time. The STS approach in contrast is computationally cheap. Only due to the high computational time, the studies presented in the following two chapters are analyzed using the STS approach taking into account that the width of the estimated parameter distributions result from the interference of parameter uncertainty and biological cell-to-cell variability.

5. Single-cell transfection studies of lipid-based mRNA nanocarriers

One of the key challenges that needs to be solved for the application of nucleic acids (NAs) as new drug entity is their efficient delivery into cells [125, 126]. For this purpose the RNA has to be complexed with some nanocarrier to enable the transfection of cells as generally described in section 2.1.. A wide range of materials is used as nanocarriers for NA transfection such as polymers, lipids, or hybrid molecules [18, 127]. The development of non-viral vectors is focusing on efficient and safe encapsulation, transport, and release of NA into cells [128, 129, 130]. The interaction with blood serum proteins, the thereby affected intracellular uptake, limited release into the cytosol and immunological response belong to the main hindrances of lipid-based vectors [14, 16]. Each of these challenges is affected independently by variations of the nanocarrier formulations and transfection conditions. RNA encapsulated in lipid vectors is taken up by various uptake mechanisms and subsequent intracellular pathways. It is reported, however, that the majority of NA delivery systems is trapped inside endosomes or lysosomes after uptake and that their release depends on endosomal recycling [131]. The uptake and fate of delivery particles has been visualized at the single-cell level by time-resolved fluorescence microscopy studies [132, 133, 134, 135, 136]. It was shown that siRNA release occurred during a narrow 'window of opportunity' demonstrating that timing and efficiency are linked [135, 137]. However, the uptake pathways and mechanisms that control endosomal release of lipid-based nanocarriers are not fully understood. In particular, the interdependence of timing and efficiency of gene delivery and the role of extracellular factors on the kinetics of cellular uptake and endosomal release are poorly understood aspects in the delivery process.

The single-cell translation assay described in section 2.3. was used like described in chapter 4 for high-throughput readout of translation kinetics. The single-cell analysis identifies the individual protein expression onset times of each translation kinetic. The assay enables to determine the distribution of expression onset and expression rate after transfection. In this context, a systematic study of nanocarrier uptake kinetics, the expression onset times, and its relation to protein expression efficiency and external transfection conditions has not yet been carried out at the single-cell level. The time-to-expression after mRNA delivery as well as the expression rate are studied using high-throughput single-cell analysis of eGFP reporter fluorescence. The single-cell trajectories are analyzed using a deterministic translation-maturation model, which expands the introduced two-stage model (Equation 4.1) by a protein maturation step. The single-cell data reveal large cell-to-cell variability and the insight that the delivery time and expression rate do not correlate at the single-cell level. Nevertheless, both parameters change in a systematic manner as a function of external parameters such as the amount of serum. In

mRNA transfection experiments using L2000 lipoplexes the efficiency decreases and delivery timing increases as a function of the fraction of serum protein. In contrast, for lipid nanoparticles (LNPs) containing the ionizable cationic lipid DLin-MC3-DMA efficiency improves and delivery time shortens in the course of increasing serum. The LNPs show a more homogeneous behavior compared to L2000 lipoplexes. As LNPs show a superior behavior compared to L2000 lipoplexes and the fact that they are one of the most promising non-viral delivery systems in current clinical trials [138], three different formulation of LNP were analyzed. The LNPs formulations distinguish only for the ionizable lipid, which have different (logarithmic) acid dissociation constant (pK_a). For siRNA transfection, the efficiency of LNPs has increased considerably during the last years and the physico-chemical properties of the siRNA-LNPs are widely studied [125, 139]. However, the use of LNPs for mRNA transfection is less well investigated. So far, it is not known how ionizable lipids of different pK_a values affect the delivery timing and expression efficiency.

In this chapter a more detailed introduction of lipid-based vectors and their uptake pathways and the determination of the onset time t_0 is given. Afterwards, the systematic variation of serum fraction during transfection and its influence on onset time and expression rate are discussed. In the following, the expression efficiency of the three LNP systems is investigated. This chapter is mainly based on the publication [P3] by Reiser et al. [140], in which the method of transfection quantification of different nanocarriers under varied conditions is presented. The study on LNP formulations with different ionizable lipids and their influence on transfection efficiency at the single-cell level is part of the manuscript [M2] listed in Associated publications and manuscripts. The affiliated data and algorithms of publication [P3] are deposited at zenodo (doi:10.5281/zenodo.2626006).

5.1. Lipid-based delivery systems and uptake pathways

Lipid-based vectors are used for the delivery of a variety of drugs like anti-cancer, anti-inflammatory, or nucleic acids and have been studied for more than five decades [141]. The fabrication and formulations of mRNA delivery systems progressed over the last 40 years from simple liposomes formed by spontaneous self-assembly to delivery systems of several components produced under standardized conditions [142, 143]. In general, lipid-based vectors are vesicles of unilamellar or multilamellar phospholipid bilayers, which surround the NA in an aqueous solution. The materials typically used are lipids consisting of polar head groups and nonpolar tails. The interaction between these groups drives vesicle formation [15]. The progress of delivery system development is driven by improving the drug loading efficiency, the release efficiency, and cell-type targeting by minimizing negative effects like immune response at the same time [141].

Lipoplexes

Lipoplexes are a commonly used lipid-based vector for NA transfection and are typically fabricated by bulk mixing of cationic liposomes, a mixture of cationic and neutral lipids,

and NA [144]. It has been shown that lipoplexes arrange as multilayer structures with NA intercalated between lipid bilayers [145]. But, lipoplex formulations are limited regarding their physical and chemical stability and the reproducibility between batches. They belong to the so called first generation of lipid-based vectors [141, 146]. In principle, each lipid-based vector can be denoted as lipoplex. In this thesis the term is used for spontaneous self-assembled complexes produced by bulk mixing of liposomes with mRNA.

In order to transfect negatively charged mRNA molecules, they need to be shielded to be capable to cross the negatively charged cell membrane, and not to be hindered by electrostatic repulsion [17]. mRNA-lipoplexes are able to fulfill this task by complexing the mRNA with cationic liposomes. The lipoplexes proved to be able to transfect many different cell types leading to a large amount of commercially available transfection agents like lipofectamine [147]. Lipofectamine was already successfully used in many *in vitro* as well as *in vivo* studies [15]. The properties of the lipoplex can be adjusted for example by different ratios of cationic to neutral lipids or the charge ratio of NA to cationic lipids. The ratio has an influence on size and surface charge of the lipoplex [148, 149]. It is commonly accepted that the properties of lipoplexes affect the transfection efficiency because these properties have an influence on different intermediate steps of the delivery process like cell membrane fusion or endosomal release [144]. The transfection efficiency is furthermore affected by external conditions like the presence of blood serum proteins or the targeted cell type [148, 150, 151, 152].

Lipid nanoparticles (LNPs)

LNPs are one of the leading non-viral vectors for gene therapy development with several active clinical trials and can be referred as lipid-based vectors of the second generation [138, 153]. Patisiran, a siRNA-LNP drug [154], was the first RNA therapeutic approved by the US Food and Drug Administration in 2018 [12]. LNP systems were originally developed for the delivery of small molecule drugs [155] and the first small molecule LNP drugs were already approved in the 1990s [153]. Most work on the development of efficient NA-LNP systems was done using siRNA. Published LNP formulation are therefore optimized for double stranded RNA with a length of $\sim$20 bp. In order to use LNPs for the delivery of single stranded mRNA molecules with typical lengths in the range of 1000 nt, which have different physico-chemical properties than siRNA, the formulations need to be optimized [156].

As it is proven that cationic lipids lead toxicity effects and immune response of the transfected host cells, ionizable lipids were developed as alternative. The positive surface charge of lipoplexes can be neutralized by the binding of negatively charged proteins in the cell environment (typically serum proteins) which leads to toxicity and decreased transfection efficiency [157]. Suitable ionizable lipids usually have a pK_a value lower than the pH at physiological conditions. If ionizable lipids are mixed with NA below their pK_a value the lipids are protonated, aiding the complexation of ionizable lipids with NA. If the pH is increased to physiological pH after complexation, the surface charge of the delivery system is neutral and reduces negative side effects, which occur for positively charged nanocarriers. The LNPs are mostly produced by rapid-mixing methods using microfluidic chips.

This manufacturing process is advantageous, in that it provides uniform properties of the complexes such as size in the range of several tens of nm, high entrapment rates of the NA, and a robust fabrication process [139, 158]. The most frequently used method for encapsulating NA in LNPs is an ethanol loading approach. The ethanol solution contains ionizable lipids, cholesterol (Chol), distearoyl phosphocholine (DSPC), and a PEG-lipid and needs to be rapidly mixed with an aqueous buffer containing the NA at low pH [153]. The hydrophobic Chol promotes the stability of the vesicle. Helper lipids like DSPC enhance the interaction of LNPs with the cell or the endosomal membrane, which is both important for uptake as well as for release. The PEG-lipid, a PEG molecule conjugated to a lipid, reduces the adsorption of many serum proteins [15]. The properties of the mentioned components and the precise formulation can have large effects on the efficiency making LNPs a very variable system.

The LNPs used in this work were prepared by cooperation partners from AstraZeneca (Sweden) by microfluidic mixing following an already published protocol based on the ethanol loading approach [159]. The LNPs contain one of the three ionizable lipids DLin-DMA (DLin), DLin-KC2-DMA (KC2), or DLin-MC3-DMA (MC3). Beside the ionizable cationic lipid, DSPC, DMPE-PEG2000, and Chol were used. The formulation of the LNPs was as follows: Ionizable lipid:DSPC:Chol:DMPE-PEG2000 in the ratio 50:10:38.5:1.5. The mRNA constructs, used for lipoplex and LNP formation, were purchased from Trilink Biotechnologies (USA). Both have identical sequences containing an ORF for eGFP (construct length: 996 nt length) but one of the constructs was additionally synthesized with a fraction of Cyanine dye 5 (Cy5) labeled nucleotides to mark the transfection complexes.

Cellular particle uptake

As already mentioned, for an efficient transfection the encapsulated mRNA need to be taken up by cells and has to be released into the cytosol. This is mainly achieved by endocytosis, an internalization process of extracellular material, solid and fluid, from the cell's environment. Endocytosis is a complex biological process in which a part of the cell membrane is inverted leading to the formation of a vesicle. The interior of the vesicle contains material from the cell environment. This way the cell membrane engulfs the adsorbed mRNA nanocarrier and transports the nanocarrier inside the pinched of endosome to the cytosol. During endocytosis the pH inside endosomes is reduced from roughly 6.8 to 4.5 by inwardly directed proton pumps, which is referred as endosome maturation. The very acidic late endosome morph with lysosomes in which trapped macromolecules are degraded [160]. To avoid the degradation of mRNA trapped in endosomes, they need to be released into the cytosol. This process is called endosomal escape. It is one of the critical barriers, which a mRNA nanocarrier needs to cross, as it has been reported that only a small fraction of nanocarriers escape successfully [134, 161]. The mechanism of endosomal escape is not well understood. But, it is proposed that cationic lipids of the delivery system fuse with endogenous anionic lipids of the endosome resulting in a change of the molecular structure of the delivery system [15]. This structural change results in a disruption of the endosomal membrane, releasing the mRNA in the cytoplasm [162]. The endosome maturation leads to a gradual protonation of the ionizable lipids enabling the

electrostatic interaction of the ionizable lipids with the anionic endosomal lipids. Studies using siRNA-LNPs showed that the release efficiency is most successful if the ionizable lipids have pK_a values in the range from 5.5 to 7 [163, 164].

As the endosomal maturation is universal, the uptake itself consist of many different pathways which can be classified based on the local membrane properties of mRNA-nanocarrier adsorption. Several studies confirmed that the cell membrane does not exist of an uniformly distributed milieu, but that region of more ordered clusters exists [165]. These clusters of 10-200 nm in diameter are named "lipid rafts" and compartmentalize cellular uptake processes [166]. Some of the most common uptake mechanism are clathrin-mediated uptake, phagocytosis, or macropinocytosis. These pathways dependent on different compositions of the lipid rafts and the fate of an adsorbed nanocarrier may already depend on the location of adsorption [165]. Pozzi et al. [144] reported, that the lipoplexes in their study where mainly taken up via macropinocytosis. In contrast, the used LNPs were taken up by clathrin-mediated endocytosis as well as macropinocytosis. These results show that lipid-based vectors can target different endocytotic pathways, which may result in different delivery timing as well as varying transfection efficiency (TE).

5.2. Single-cell protein expression readout by on stage transfection

The single-cell microarray approach described in section 2.3. was used to obtain hundreds of single-cell translation kinetics of mRNA-L2000 transfected HuH7 cells with focus on the protein onset time and expression rate distributions. The single-cell traces show large heterogeneity regarding the expression levels like previously described [21, 33, 83] and a large variability for the onset time points like shown in Figure 5.1..

For estimation of the onset time distribution without loss of the early onset time points the use of a perfusion system is essential otherwise the transfection need to be done off the microscope. But the transfection off the microscope makes it impossible to measure changes in fluorescence intensity during the incubation step and the set up of the time-lapse measurement, which last at least 1.5 h. Furthermore, the use of a perfusion system has the advantage of a pulse like transfection incubation of 1 h period (gray bar in Figure 5.1.) during the time-lapse measurement that enables to control the possible time for particle uptake. The short incubation time was chosen to eliminate the possibility of late mRNA particle uptake, which cannot be excluded without the flushing step of the unbound particles. The complex adsorption kinetics and removal of unbound L2000 lipoplexes, which contain Cy5 labeled mRNA, are shown and discussed in Figure 7.6.. The washing steps are indicated as black arrows and after the incubation period the medium is exchanged to cell growth medium for the remaining measurement. The magnification in Figure 5.1. illustrates the variability for the onset time points between different cells and that some cells already started to express protein during complex incubation or shortly afterwards. These time points would be not detectable without the on stage transfection.

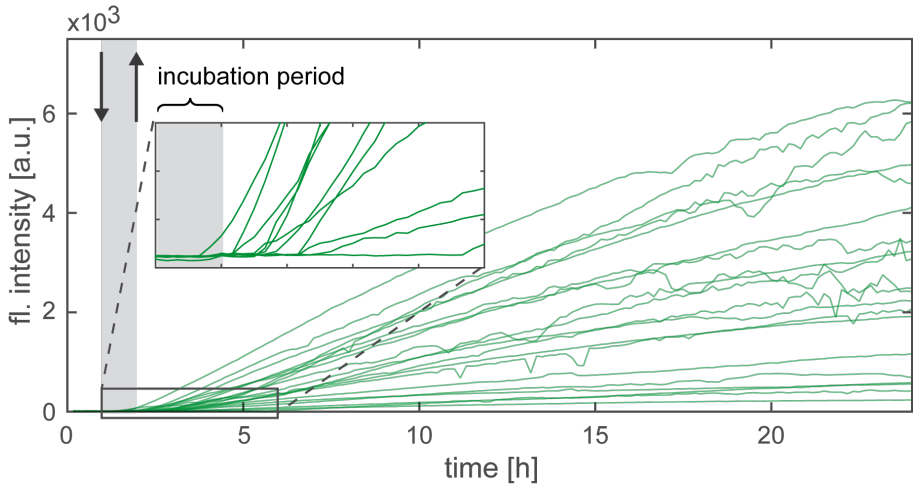

Figure 5.1.: eGFP expression trajectories (green) show large cell-to-cell variability after mRNA transfection. The pulse like transfection of 1 h (gray bar) is performed by flushing a connected perfusion system for fluid exchange of the microarray channels (exchange steps are indicated by the arrows). The transfection on stage enables to record all expression onset time points, even the early events, within the incubation time like illustrated in the enlarged plot.

5.3. Onset time determination

The onset time of protein expression need to be determined automatically to allow for high-throughput and reproducible measurements. For reliable detection of the onset time point for each single-cell trajectory, three different methods were tested and are explained and compared in the following subsections.

5.3.1. Translation-maturation model

Fitting a translation-maturation model to the single-cell trajectories was found to be the most reliable method. With the fitting approach the kinetic rates and the protein expression onset time were obtained. The three-stage translation-maturation model shown in Figure 5.2. A is based on biochemical rate equations like the translation models discussed earlier in chapter 4.

$$\begin{aligned} \frac{d[mRNA]}{dt} &= -\delta[mRNA],\ [mRNA](t_0) = m_0 \\ \frac{d[GFP^*]}{dt} &= m_0 k_{TL} - k_M[GFP^*] - \beta[GFP^*],\ [GFP^*](t_0) = 0 \\ \frac{d[GFP]}{dt} &= k_M[GFP^*] - \beta[GFP],\ [GFP](t_0) = 0 \end{aligned} \tag{5.1}$$

The model considers a translation step to a non fluorescent state of the reporter protein GFP^* from the transfected $mRNA$ with rate k_{TL}. The translation is followed by a matu-

ration step [51, 167], which leads to a chemical transition of non fluorescent GFP^* to the fluorescent protein GFP with maturation rate k_M. Thereby, mRNA and protein undergo degradation with the rate δ and β, respectively, with the assumption that the two protein states have the some degradation rate β. The same protein degradation rate was assumed because the unmatured and matured GFP is degraded by the same proteases [168]. The three ODEs in Equation 5.1 describe the temporal concentration changes of $mRNA$, unmatured protein GFP^*, and matured protein GFP. The analytical solution for GFP of the above ODEs is as follows:

$$GFP(t) = m_0 k_{TL}\left(\frac{k_M}{(\delta-\beta)(\delta-\beta-k_M)}e^{-\delta(t-t_0)} - \frac{1}{\delta-\beta-k_M}e^{-(\beta+k_M)(t-t_0)} + \frac{1}{\delta-\beta}e^{-\beta(t-t_0)}\right) \quad (5.2)$$

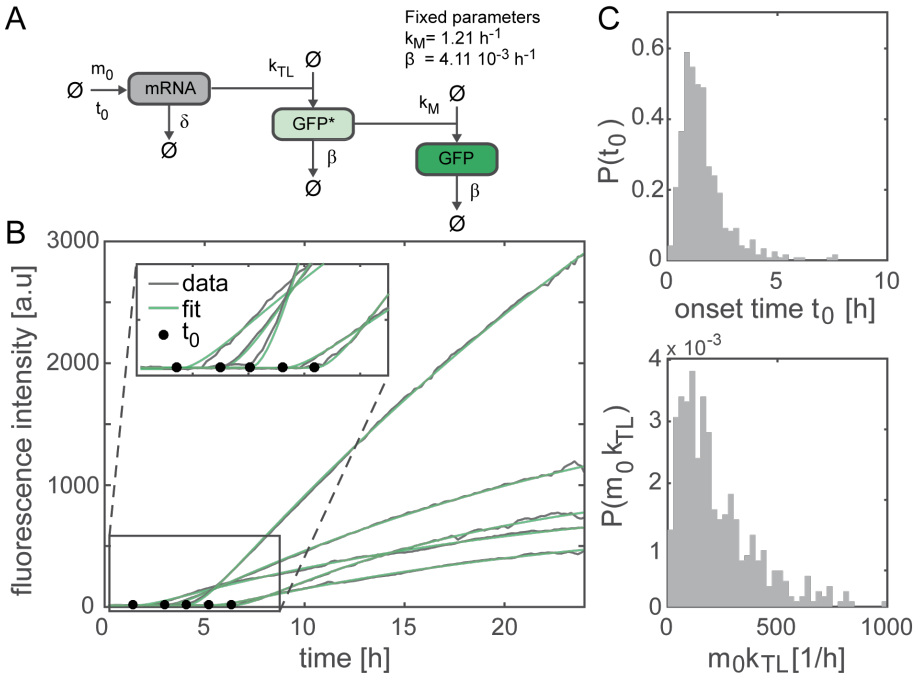

Figure 5.2.: Determination of onset time point t_0 and the expression rate $m_0 k_{TL}$ by fitting the translation-maturation model. (A) The biochemical network of translation followed by protein maturation is sketched with the corresponding parameters next to the reaction. The number of parameters is reduced by fixing the maturation rate k_M and protein degradation rate β to population mean values. (B) The exemplary traces (gray) with the respective fits (green) show the good agreement of the fit with the model. The onset times are indicated as black dots. (C) The histograms show the distributions of the two parameters of interest.

Here, m_0 is the number of successfully transfected mRNA molecules at the onset time point t_0. The delivery time is defined as the duration between transfection start and

protein expression onset. All determined onset times are relative to the time point of lipoplex addition to the cells. The delivery time includes therefore all intermediate steps like complex uptake or mRNA release. The model assumes one mRNA release event for all molecules, which is clearly a simplification. However, the duration of mRNA release is short compared to the total expression and is therefore treated as one single event. The product of released mRNA with the translation rate $m_0 k_{TL}$ is fitted as one parameter due to parameter identifiability and a fluorescence offset parameter z was added. A detailed description of the fitting routine with the corresponding starting values is given in the Material and Methods section of the publication [P3]. The analytic solution (Equation 5.2) was fitted to each single-cell trajectory using a Least-Square approach, which correspond to the maximum likelihood fitting described in section 2.4 for the assumption of Gauss-distributed measurement noise.

To robustly determine the parameters t_0 and $m_0 k_{TL}$ the parameters k_M and β were fixed to mean values to avoid over fitting. The protein maturation and degradation rate were determined by a translation block experiment using cycloheximide, which was performed as an independent experiment previously [119] and the protocol is described in Appendix A. With fixing the maturation and degradation rate the number of free parameters are reduced to four. The parameter set, which is determined for each cell, consists thereby of the fluorescence offset, the mRNA degradation rate, the expression rate, and the onset time. Figure 5.2. B illustrates the good agreement of the fit with the experimental data for representative single-cell trajectories. The reliable determination of the onset times (black dots) is illustrated in the magnification of the example trajectories in Figure 5.2. B.

In order to estimate the population distributions for the parameters of interest, the onset time and expression rate, the translation-maturation model was fitted to each single-cell trajectory (Figure 5.2. C). The onset time distribution shows a width of roughly 2 h for the slightly skewed distribution. The earliest onset times occur only a few minutes after adding the transfection particles, while the later onset times extend to several hours. The expression rates show a broad distribution over three orders of magnitude.

5.3.2. Alternative approaches for onset time determination

To enable high-throughput and reproducible results, three approaches were compared for reliable onset time determination. The described translation-maturation model appeared to be the most suitable approach. Besides this model, the analytical solution of the two-stage translation model (Equation 4.3) and a model independent approach based on hierarchical clustering analysis were tested.

The clustering approach is explained in detail in the supplementary information of the publication [P3]. For the clustering approach, the onset time determination is treated as a change point detection problem [169]. The change point is defined as the border between the time point clusters before and after the onset time. The fluorescence intensities fluctuate around an offset value z before the onset time. These fluctuations are small compared to the eGFP fluorescence signal, which increases for several hours after the onset time. This behavior leads to a densely populated interval of fluorescence values

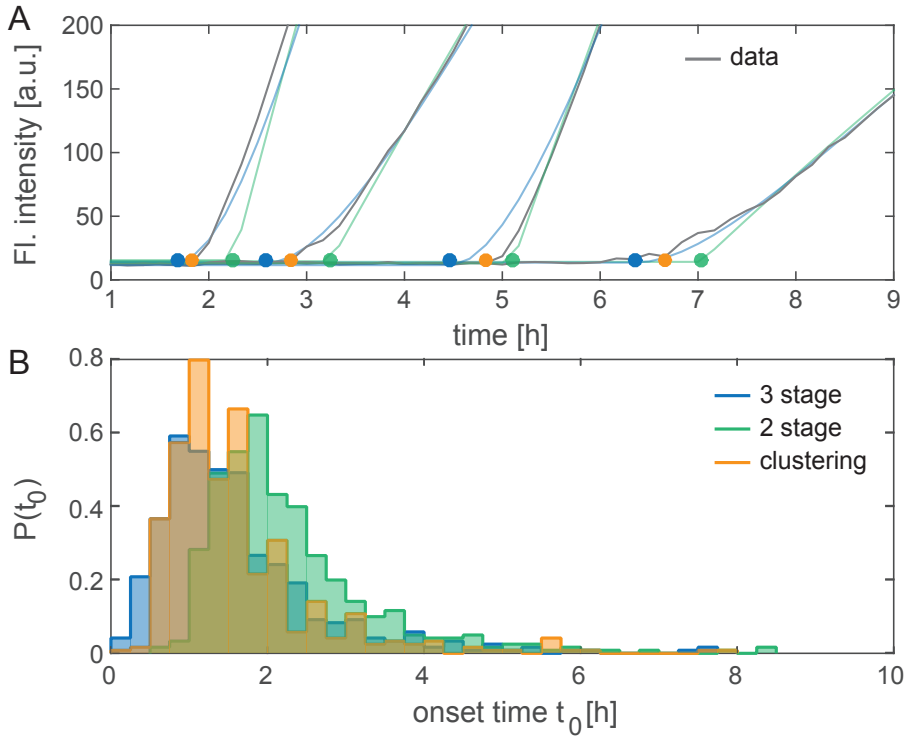

Figure 5.3.: The three methods show differences regarding the determination of the onset time t_0. (A) Exemplary trajectories (gray) are shown with the respective fits for the two-stage (green) and the three-stage (blue) model. The determinated onset times are shown as dots with the corresponding color together with the onset time determined by clustering (orange). (B) The distribution of the two-stage onset times is shifted to longer times compared to the other two approaches.

before the onset time and a population of lower density afterwards. The differences of the densities before and after translation onset is used to determine the onset time point by hierarchical clustering. To determine the border between the clusters a cutoff value need to be defined manually.

Figure 5.3. illustrates the differences of onset time determination for the three methods. Figure 5.3. A shows representative trajectories (gray) with the respective fits for the two-stage (green) and the three-stage (blue) model. The two-stage model implements the onset as a sharp kink, which leads generally to an overestimation of the onset time. This overestimation is represented in the high residuals between fits and experimental data like discussed in Figure 4.5. D. In contrast the translation-maturation model represent the same onset behavior than the data. The clustering onset times (orange) are typically located between the two model approaches and tend to be closer to the three-stage model. These differences are also reflected in the distribution of the onset times shown in the histograms of Figure 5.3. B. The histogram of the two-stage model is shifted to later times compared to the other two approaches. The cluster approach leads to good results if a suitable cutoff value is set to determine the border between the two clusters. This value can vary dependent on the data set and has to be manually chosen, which is a

disadvantage compared to the three-stage model.

5.4. Transfection ability changes dependent on external conditions

In order to investigate if the delivery time is correlated with the expression efficiency at the single-cell level, fluorescence trajectories were analyzed using the translation-maturation model obtained from cells transfected with different lipd-based vectors or varied FBS fractions. As a model system mRNA-L2000 lipoplexes were used to study the effect of nonspecific serum protein binding on the nanocarrier's surface by increasing the FBS fraction within the medium during the 1 h of complex incubation. To exclude day-to-day variances, six FBS fractions ranging from 0% to 10% (v/v) FBS were investigated in parallel. Each FBS fraction was applied in one of the six channels of the single-cell array with approximately the same number of cells per channel.

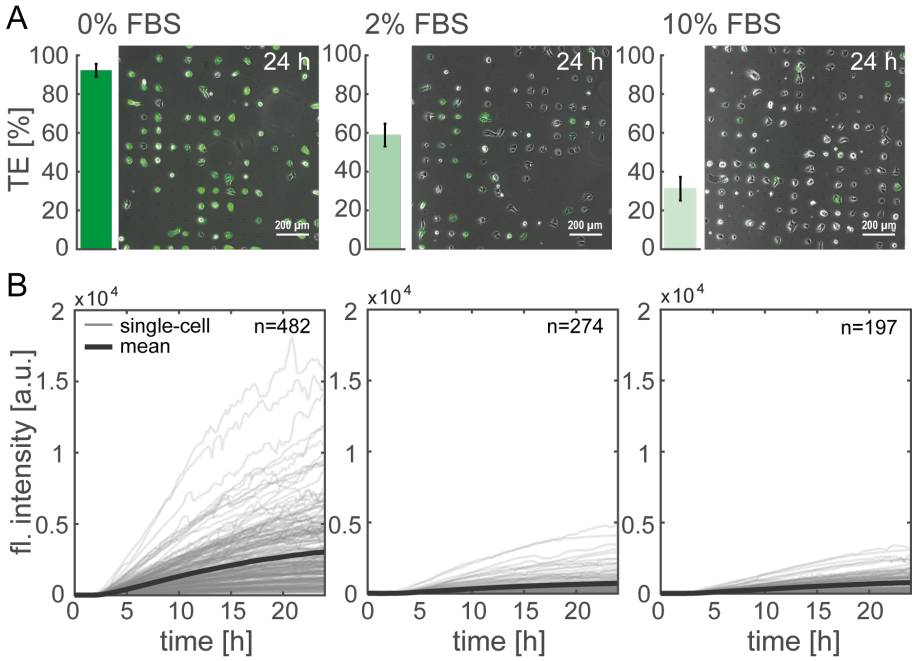

Figure 5.4.: The protein expression kinetics show changes dependent on FBS concentration reflected in TE and the single-cell time courses. (A) The TE (mean ±std) decreases with increasing FBS fraction. Representative overlays of the phase contrast image with the eGFP fluorescence image are shown with constant contrast settings for direct comparison of intensities. (B) The single-cell trajectories show the eGFP expression kinetics (gray lines) and the respective mean trajectory (black line) of each data set with the number n of cells. The data sets of cells treated with 0%, 2%, and 10% FBS are exemplary shown.

The representative overlays of the phase contrast images with the corresponding eGFP fluorescence images at the end of the measurement, 24 h after transfection, show clear

differences in the fluorescence intensities as well as in the number of fluorescent cells for different FBS fractions (Figure 5.4. A). The differences in TE, defined as the ratio between eGFP positive cells and the total number of cells per field of view, is determined using all endpoint images. The TE decreases from 92% ±3% for cells treated without FBS to 31% ±6% for cells transfected with 10% FBS (the mean TE data for 0% FBS, 2% FBS, and 10% FBS is shown with the respective standard deviation in Figure 5.4. A). The decrease of TE in the presence of serum proteins is known from literature [152, 170]. The single-cell trajectories of the three data sets with the respective mean trajectories are shown in Figure 5.4. B. A large difference of the expression kinetics is observed between cells treated without FBS to cells treated with FBS, which is in agreement with the decreasing TE. The serum's influence is reflected in the decreasing number of single-cell trajectories, which is expected due to the decreasing TE under the assumption of the same number of cells per condition.

5.4.1. eGFP expression kinetics change with increasing serum fraction

The fluorescence trajectories of the six FBS data sets are analyzed by fitting the translation-maturation model as described in subsection 5.3.1. to obtain the parameter set of the four free parameters for each cell. The fluorescent offset z, which has no biological meaning, and the mRNA degradation rate δ show no dependence on varying FBS concentration. Figure 5.5. shows the histograms for the fluorescence offset z (A) and the mRNA degradation rate δ (B) with the respective mean values (black circles). Both parameter show only statistical fluctuations for all data sets. The background offset has a mean value of 13.5 ±0.4 a.u. and the mRNA degradation rate is 0.052 ±0.006 h^{-1}. The mean mRNA degradation rate equates to a mRNA half-life of approximately 13.4 h of the transfected construct.

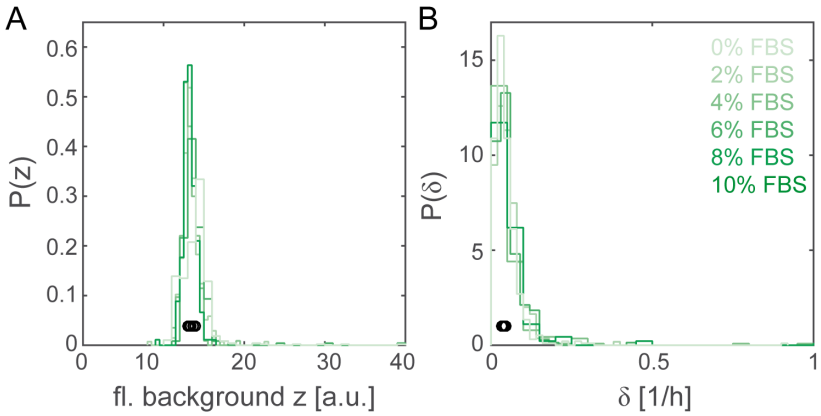

Figure 5.5.: The histograms (green) of the fluorescence offset z (A) and the mRNA degradation rate δ (B) show non significant variations between the single data sets of different FBS fractions, which is further illustrated by the overlapping mean values (black circles) of each distribution.

The other two free parameters, t_0 and $m_0 k_{TL}$, of the translation-maturation model show systematic changes for changing FBS levels. Figure 5.6. A shows the parameters for t_0 and $m_0 k_{TL}$, whereas each dot correspond to the parameter set of one cell. The mean value and the standard deviation as error bar illustrate the cell-to-cell variability for each data set of both parameters. The onset time distributions shift to later times with a shift of the mean onset time from 1.6 ±1.3 h for 0% FBS to 2.9 ±2.1 h for 10% FBS.

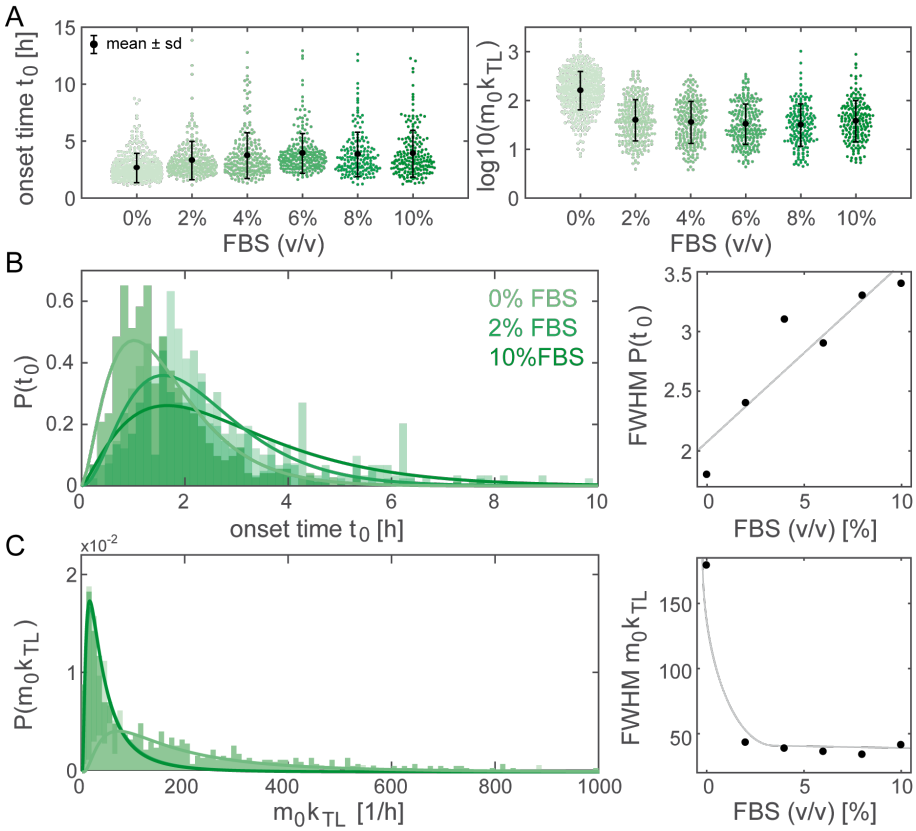

Figure 5.6.: (A) The single-cell parameters for the onset time and the expression rate (green dots) change dependent on the FBS fraction. The mean value with respective standard deviation is shown in black. (B) The onset time shifts to later times with increasing FBS level. The distribution can be described by Gamma functions respectively shown for 0%, 2%, and 10% FBS. The distributions broadens with increasing FBS measured by FWHM of the Gamma distributions. (C) The expression rate distributions gets smaller if FBS is present. Log-normal distributions are used to fit the histograms and to calculate the FWHM. The gray lines of the FBS vs. FWHM plots are to guide the eye.

The shift of later onset times is co-occurring with a broading of the distributions, which is measured by the full width half maximum (FWHM) of the distribution (Figure 5.6. B). The skewed distributions of the onset time histograms are thereby well described as Gamma-distributions. The FWHM shows a linear increase from 1.8 h to 3.4 h with increasing FBS fraction. The shift of the distributions reveal a delayed mRNA release into

the cytosol with increasing variance between the data sets regarding higher FBS fractions, which can be seen by the increasing FWHM of the onset distributions. The expression rate seems to be more sensitive on increasing FBS than the onset time, which can be seen by the break-in of the expression rate to almost a steady state level for all cells treated with FBS. The expression rate distribution widths of FBS treated cells get much sharper compared to the cells without FBS (Figure 5.6. C). The expression rate decreases by factor of four between untreated cells and cells treated with 10% FBS (Figure 5.6. C). The distributions for all six data sets is thereby well described by log-normal distributions like shown in Figure 5.6. C. The decrease of expression rate could mainly result from the decreased number of released mRNA m_0. It is expected that the translation rate k_{TL} is not affected.

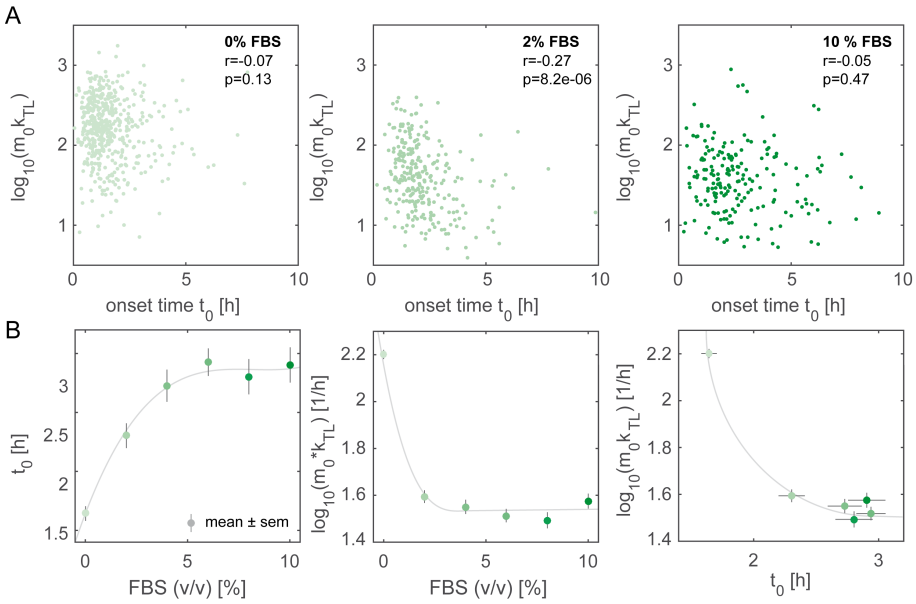

Figure 5.7.: (A) Scatter plots of the single-cell data of the onset time vs. the expression rate do not show correlation. The Pearson's correlation coefficients r are shown with the respective p-values for three data sets of different FBS levels. (B) Both parameter systematically change with increasing FBS fractions (left and middle plot) resulting in a negative trend of t_0 vs. $m_0 k_{TL}$ for increasing FBS level (right plot). The population mean values are plotted with the standard error of the mean (sem) with color intensity indicating the FBS content. The gray lines are to guide the eye.

The single-cell onset times and expression rates were then further analyzed to investigate if fast delivery is linked with high expression efficiency. The scatter plots of the onset times with the expression rates are uncorrelated at the single-cell level for all six data sets, which is exemplary shown for 0%, 2%, and 10% FBS in Figure 5.7. A. The fact that the onset time and the expression rate is uncorrelated at the single-cell level is indicated by the Pearson's correlation coefficients r, which range between -0.07 and -0.27 for all data sets. The low coefficients indicate that there is no intrinsic mechanism linking the total delivery time with the expression efficiency. But, the parameters show a systematic change

on the population level (see Figure 5.6. B). The mean onset time increases until 4-6% FBS and show only small changes within the error for higher FBS fractions (Figure 5.7. B first plot). The mean expression rates decrease and show no significant changes if FBS is present (Figure 5.7. B second plot). The population behavior of both parameters lead to negative trend for increasing FBS fractions (Figure 5.7. last plot).

5.4.2. Effect of protein adsorption differs between lipoplexes and LNPs

The addition of serum proteins lead to a decrease of the expression rate with an increase of the onset time for L2000 lipoplexes. However, this dependency is not universal for lipid-based vectors. If single-cell data of cells transfected with L2000 or LNPs is compared an opposite behavior is observed (Figure 5.8.). LNP transfected cells show an increase of expression rate and a shift to earlier onsets if FBS is present during transfection. The cells treated with the ionizable MC3 LNPs show less cell-to-cell variability for both parameters compared to lipoplex transfected cells, which can be seen in the less scattered data (Figure 5.8.). The onset time of LNPs shifts from 4.4 ±0.7 h for untreated cells to

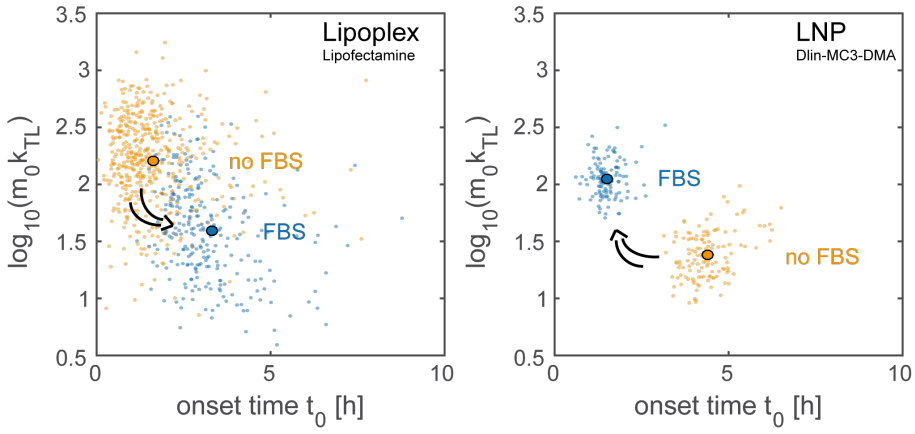

Figure 5.8.: Protein adsorption on the mRNA vector's surface has the opposite effect on LNPs compared to lipoplexes. The single-cell scatter plots for the onset time against the expression rate reveal the opposed effect on transfection ability between L2000 lipoplexes (left) and LNPs (right) dependent on FBS addition. The arrows clarify the opposite effect induced by FBS. Each data point correspond to the parameter set of a single cell and the median value for each data set is shown as thick dot.

1.5 ±0.4 h of cells transfected with FBS in the present. For comparison the onset time of L2000 transfected cells shift from 1.6 ±1.3 h to 2.3 ±1.7 h (2% FBS data set). The smaller cell-to-cell variability is also indicated by the smaller standard deviations of less than one hour for the LNP treated cells. The mean expression rate increases by a factor of four from 30 ±15 h^{-1} to 120 ±50 h^{-1} after FBS addition. But, like for lipoplexes the single-cell parameters are uncorrelated at the single-cell level for both cases.

5.5. Ionizable lipids and their effect on expression efficiency

The MC3 LNPs show a more suitable behavior compared to L2000 lipoplexes with regard to therapeutic purposes. The cell-to-cell variance is less and they are functional if serum proteins are present, which is crucial for *in vivo* applications. In order to investigate how ionizable lipids with different pK_a values influence the expression efficiency, three different LNP formulations introduced in section 5.1. were tested. The analyzed LNPs are made with one of the ionizable lipids: DLin (pK_a=6.8 [171]), KC2 (pK_a=6.7 [171]), or MC3 (pK_a=6.4 [172]). The three LNP formulations were analyzed using the same assay as for the L2000 lipoplex data sets. The quantification of the expression efficiency for the different LNPs are part of the manuscript [M2] listed in Associated publications and manuscripts. In this work, the structural changes of the LNPs under varied pH values measured by X-ray scattering are compared to their transfection ability on a population and single-cell level. The goal is to obtain a better understanding how internal LNP structure variation due to pH changes, which occurs during endocytosis, effect the transfection efficiency of the delivery system.

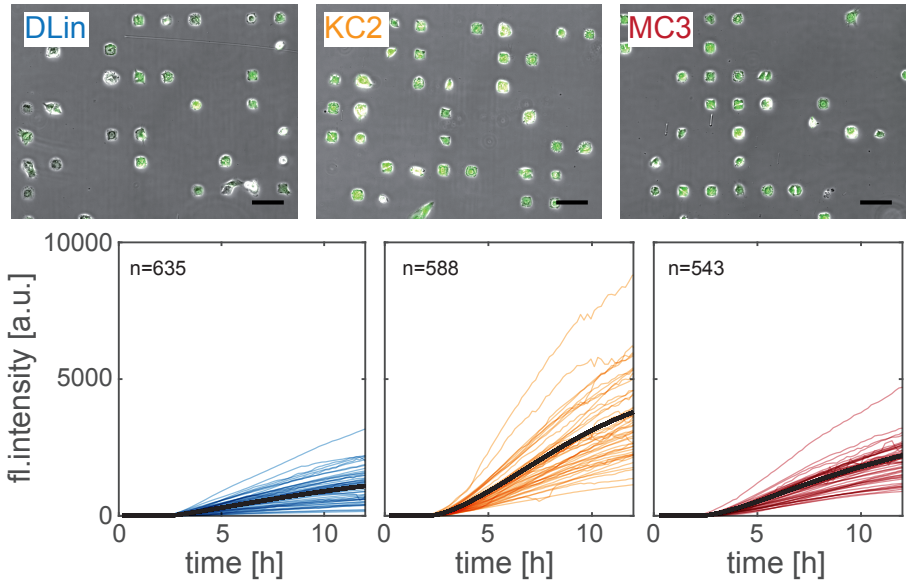

Figure 5.9.: : Transfection of cells using different mRNA-LNPs. Cells expressing eGFP (green) on a micropatterned array transfected with one of the used LNPs 10 h after transfection. The scale bar correspond to 100 µm. The single-cell trajectories show the eGFP translation dynamics of cells transfected with DLin (blue), KC2 (orange), or MC3 (red). The black time course correspond to the respective population time course.

HuH7 cells were transfected with one of the LNP formulations 1 h after the time-lapse measurement was started with a final concentration of 0.5 ng/µl eGFP mRNA in L15 medium supplemented with 1% FBS. The LNPs were incubated for 1 h like in the previous experiments. Figure 5.9. shows a subset of the single-cell trajectories for the three LNP

formulations. Comparing the translation kinetics, it is obvious that the expression levels differ between the LNP systems. The KC2 particles show the highest expression, followed by cells transfected with MC3. DLin shows the weakest expression. Interestingly, all three LNPs seem to have a narrow time interval for expression onset compared to trajectories transfected with L2000 (Figure 5.1.).

This impression is confirmed by the similar distributions of the onset times (Figure 5.10. A). All LNPs show a mean onset time of ∼1.4 h with small variation regarding the standard deviation of 0.2 h for MC3, 0.3 h for KC2, and 0.4 h for DLin. The same trend is observed for the FWHM of Gamma distributions fitted to the onset times (MC3 0.5 h, KC2 0.6 h, DLin 0.8 h). The expression rate on the other hand clearly changes for different ionizable lipids (Figure 5.10. B). The distribution for DLin has the smallest mean value of 161 $\pm 95\,h^{-1}$ followed by MC3 with a mean of 310 $\pm 129\,h^{-1}$. The expression rates for KC2 is the highest with 538 $\pm 269\,h^{-1}$. The FWHMs of the log-normal distributions fitted to the histograms illustrate that the distribution width of KC2 ($FWHM_{KC2}$=455.3 h^{-1}) is larger than the FWHM of 147.1 h^{-1} for DLin and 239.3 h^{-1} for MC3. The translation rate k_{TL} is expected to be independent on the ionizable lipid, which means that the number of successfully released mRNA molecules m_0 needs to change depending on the ionizable lipid. The scatter plot stresses the similar onset time distributions and the variation for the expression rate distributions (Figure 5.10. C). The two parameters do not correlate at the single-cell level for cells transfected with KC2 or MC3 like indicated by the small Pearson's correlation coefficients. Interestingly, cells transfected with DLin show a weak negative correlation of r=-0.33.

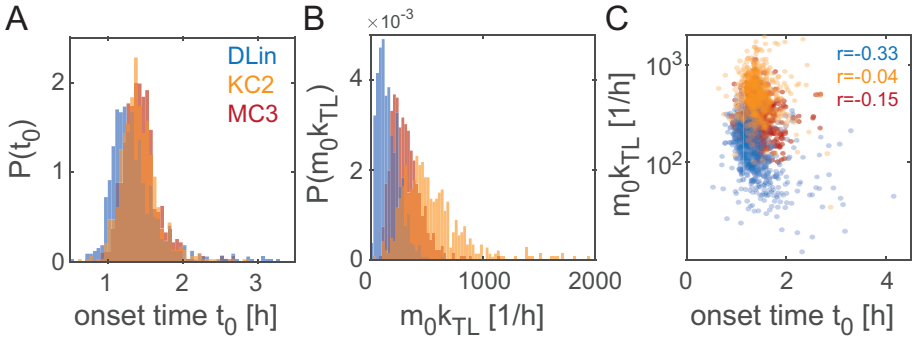

Figure 5.10.: The onset time distribution (A) for all LNP formulations do not deviate much from each other, whereas the expression rate shows clear changes (B). The scatter plot of both parameters illustrate the shift for the expression rates (C). Each data point corresponds to the parameters of one cell and the Pearson's correlation coefficients r are given for each LNP system.

5.6. Discussion

The single-cell translation assay enables measuring the expression onset and rate in individual cells in a highly parallel manner. The onset time is thereby a measure for the

mRNA transport process consisting of the intermediate steps of particle adsorption, uptake, intracellular trafficking, endosomal release, and translation. To avoid long-term nanocarrier incubation and the ability to measure early protein expression onsets the use of a perfusion system is necessary. Furthermore, the expression onset time relative to the transfection start only corresponds to the total delivery time if the incubation time is kept short. For long-term lipoplex incubation, the distribution shifts to later times and broadens compared to the onset times of a defined incubation period of 1 h like illustrated in Figure 5.11., which verifies the need of short nanocarrier incubation periods.

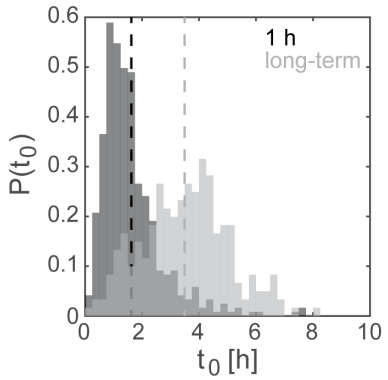

Figure 5.11.: The onset time distribution shift to later times for long-term lipoplex incubation compared to a controlled incubation period of 1 h. The mean value for each distribution is shown as dashed line.

The short lipoplex incubation enables the measurement of mRNA delivery time in individual cells and its correlation with expression efficiency. The expression rate was used as a measure for transfection efficiency at the single-cell level. It is strongly correlated with the fluorescent intensity at a distinct time point after transfection, which is often used in literature [173, 174]. The correlation can be explained because the slope of a straight line between the expression onset and a certain protein level after a certain time interval (of for example 12 h) correspond to the translation rate k_{TL} (Figure 5.12.). This approximation can only be made as long as the protein increase is in a linear regime.

The distribution of onset times with single-cell resolution was measured for the first time. They generally show narrow width of only a few hours (Figure 5.6. B and Figure 5.10. A). It is reported in literature that there is only a small window of opportunity for successful molecule release before the nanocarrier is trapped in endosomes or lysosomes [135, 137]. Since, we restrict the lipoplex incubation time to only 1 h the onset times correspond to the delivery time distribution. If mRNA nanocarriers get stuck in endosomes or lysosomes and are possibly degraded, the release of functional mRNA molecules for late time points is less likely and explains why late onsets are not observed. This mechanism would suggest that fast delivery time is linked with high molecule release and therefore with high expression rates. Yet, an general intrinsic link for these parameters could not be confirmed at the single-cell level due to the low Pearson's correlation coefficients for L2000 lipoplexes (Figure 5.7.) as well as for LNPs (Figure 5.10.). However, external parameters like increasing serum protein (FBS) concentrations can lead to a systematic change of

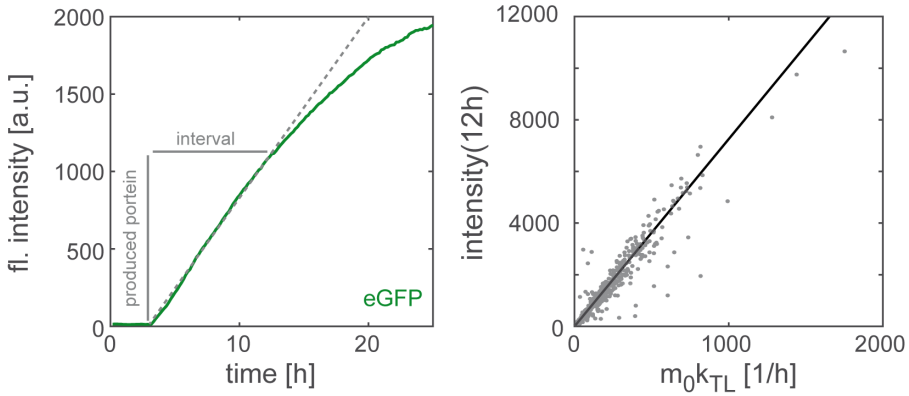

Figure 5.12.: The expression rate $m_0 k_{TL}$ shows strong correlation with the intensity level 12 h after transfection because the translation rate is approximately the straight's slope between the onset time and the fluorescence intensity after a 12 h time interval (left plot). The scatter plot of the expression rates, determined by fitting the three-stage model, and the fluorescence intensities at t=12 h as measure for transfection efficiency TE reveals the strong correlation (right plot).

both parameters. Interestingly, the effect of serum proteins on the delivery system can have the opposite effect for different nanocarrier systems. In case of L2000 lipoplexes the onset distribution shifts to later times and the expression rate decreases if FBS is present during transfection. Whereas, the opposite is the case for cells transfected with MC3 LNPs (Figure 5.8.). The unspecific adsorption of serum proteins can lead to the formation of a protein corona on the nanocarrier's surface [175, 176]. The surface charge and internal structure of the delivery particles changes depending on the protein corona formation, which influences the particle delivery for example during adsorption or endosomal uptake [177]. The formation of a protein corona appears to have different effects on the delivery time and transfection efficiency. One explanation may be the different targeting of endocytotic pathways due to different properties of the delivery systems like the lipid composition or size [144]. Proteomic studies of the corona revealed that the proteins forming the corona varies regarding their composition and quantities for different nanocarrier systems [178], which differs for lipoplexes and LNPs. The reason why LNPs show faster and more efficient uptake might be attributed to apolipoprotein E (apoE) adsorption, which is one of the proteins contained in FBS [179]. It is reported that apoE adsorption on LNPs containing ionizable lipids and dissociable PEG-lipid increases the uptake efficiency in hepatocytes [153]. Typically, the LNPs have a neutral surface charge under physiological pH, which is the reason for the poor uptake of LNPs without serum proteins. But with serum proteins present during transfection, the adsorption of apoE leads to a targeting effect via apoE recognizing receptors [180]. The reason why LNPs without FBS are still able to transfect HuH7 cells, although rather slowly, might be because HuH7 cells secret apoE leading to small amount of apoE present in the cells' environment compared to medium supplemented with FBS [181]. In summary, the effect of serum protein addition is adverse on both timing and efficiency for L2000 lipoplexes, while, in contrast, it improves both timing and efficiency for lipid nanoparticle mediated

transfection.

As LNPs are currently the most mature and clinical advanced delivery system [138], three LNP systems with ionizable lipids of different pK_a were further analyzed. All three LNP systems show similar delivery time distributions (Figure 5.10. A). The similarity regarding the onset timing is not surprising due to the same biological identity of all LNP systems. It is proposed in literature that the protein corona of a nanoparticle interacts with the cell's surface and thereby defines the biological identity of the nanoparticle [175]. For the process of LNP adsorption and the particle uptake, it is expected that the dominating properties should be the apoE containing protein corona on the PEG lipid shell of the particles [182]. The common properties of the LNP's outer shell could contribute to the comparable delivery time distribution. Besides the comparable outer shell composition, the ionizable lipids should show the same behavior at physiological pH, which is higher than all three pK_a values. These two reasons could explain why the delivery time distributions show no considerable dependency regarding the ionizable lipid. However, the expression rate distributions show a clearly different behavior (Figure 5.10. B). It is expected, that the inner structure's properties of a delivery system make the pivotal difference after uptake and influence the endosomal release efficiency [177]. As the pH declines during endosome maturation the LNP systems show different behavior due to structural changes of the LNPs driven by the properties of the ionizable lipid [183]. Therefore, it is not surprising that the expression rates are different between the LNP systems. The pH dependent changes should have an influence on the fusogenic properties of the LNPs with the endosomal membrane leading to varied amounts of released mRNAs. Comparing the three LNP systems regarding their transfection efficiency, it was observed that MC3 leads to the most homogeneous behavior taking the distribution widths of delivery time and expression rate together. KC2 shows especially for the expression rate a higher cell-to-cell variability. DLin shows the weakest transfection efficiency with the broadest onset time distributions (Figure 5.10. C). Comparing the three LNP data sets, it was observed that a change of the population distributions of the expression rate not necessarily needs to have an effect on the onset time distribution as it was the case for the addition of FBS compared to L2000 lipoplexes and MC3 LNPs (Figure 5.8.). As for the L2000 single-cell data sets, the LNP data sets show at most a weak correlation for delivery time and expression rate at the single-cell level reinforcing the assumption that there is no general intrinsic mechanism linking fast delivery time with high expression efficiency.

For further investigations of the approach, if one of the intermediate steps of the delivery process is crucial for efficient mRNA release and maybe linked intrinsically, improvements are needed. It was shown in other studies using scanning-time-lapse microscopy that more than one fluorescent marker can be used to correlate different fluorescent signals for individual cells [75, 184]. An additional fluorescent marker for labeling the nanocarriers or endosomes might be useful to gain a deeper understanding on the delivery process. For this purpose Cy5 labeled mRNA was encapsulated in L2000 lipoplexes and LNPs to observe the nanocarrier signal kinetics (Figure 7.8.). Furthermore a fluorescent dye, which labels acidic organelles such as mature endosomes or lysosomes called Lysotracker™ (Thermo Fisher), was tested for its capability to monitor endosomal uptake dynamics. Both approaches are discussed in detail in chapter 7. The labeled L2000 lipoplexes show

the lipoplex adsorption on the cell surface, but do not give further insights into the intermediate steps of transfection (Figure 7.6. B). This finding is in agreement with a previous study, where it was shown that the lipoplex adsorption kinetics for transfected and non transfected cells are the same [174]. Establishing a fluorescent dye that resolve the time event of an intermediate step of delivery, need to be carefully tested to enable event-time correlation.

To conclude, the single-cell translation assay enables to determine the onset and expression rate distributions depending on a variation of external conditions as well as for different nanocarrier systems. It offers insights into single-cell correlations of delivery timing and expression efficiency, which gives a new opportunity to quantify transfection efficiency.

6. mRNA half-life as function of poly(A) tail length

In order to use the full potential of mRNA as a therapeutic, a better understanding of how the translation process and therefore the amount of protein can be controlled is needed. In chapter 5 it was shown that the choice of mRNA transport system and the environmental conditions can have a tremendous effect on the delivery time and the expression rate. However, it is expected that the kinetic rates describing the translation process, namely the translation and degradation rates, are unaffected by the transport system. This was shown for the degradation rates in the previous chapter (Figure 5.5. B). The expression rate distributions are expected to be changed due to a varied number of transfected molecules and not due to a varied translation rate. In order to change the expression kinetics under the same transfection kinetics, the kinetic rates need to be varied. These rates can be altered by systematically varying the sequence of the exogenous mRNAs. As shown in chapter 4, the change of the protein degradation rate had an obvious effect on single-cell translation kinetics. The increase of the protein degradation rate beween eGFP and d2eGFP was obtained by adding the codon sequence of four additional amino acids to the ORF coding for eGFP resulting in the expression of d2eGFP. The variation of the protein degradation rate may be critical for therapeutic purposes, as it is possible that other protein functions beside the stability are affected as well. While a change in mRNA degradation rate is not expected to affect the protein function, it is one of the key parameters for post-transcriptional expression regulation [43]. Much effort was put into reducing the degradation rate and thus increasing the half-life of IVT mRNA. To this end, analogs for the 5' cap were engineered [47], modified nucleotides were used for synthesizes [185], the influence of UTRs and their combinatory effect on mRNA stability were studied [33, 186], and the effect of poly(A) tail length was investigated [187, 188]. These studies showed that the fate of exogenous mRNA in living cells is highly dependent on its sequence as it influences translation efficiency as well as mRNA half-life. However, it is not well understood how mRNA sequence modifications change mRNA stability at the single-cell level. Not to mention how these changes can be functionally described.

In a previous study, Ferizi et al. showed that the mRNA degradation rate distributions of IVT mRNAs change significantly for different UTR combinations using a LISCA setup similar to the studies of this thesis [33]. They showed that the approach is able to determine the functional mRNA half-life in living cells. This is an advantage compared to several other techniques used for mRNA half-life determination that measure the physical half-life of mRNAs, additionally most of these studies are performed in yeast [189]. For protein expression kinetics it is crucial that the mRNA is functional which means that it is still translatable and not only physically present. However, it is difficult to understand

the mechanistic reasons why mRNAs with certain UTR combination are more stable than others. In order to gain a better understanding of how the mRNA degradation rate can be systematically varied, the dependency of poly(A) tail length on mRNA stability was investigated. It is known for decades that the poly(A) tail is important for translation efficiency and stability regulation [38]. Yet, the effect of varied poly(A) tail length on protein expression kinetics is not well explored on a quantitative level. For this purpose, ten different mRNA constructs of poly(A) tail lengths within the physiological range for mammalian cells were investigated using the single-cell translation assay (section 2.3.). The common design for all constructs consist of the minimum components of a mature mRNA necessary for protein expression including a 5' cap, the mRNA body containing the ORF, and the 3' poly(A) tail.

Hundreds of single-cell kinetics were measured for each construct of different poly(A) tail length to estimate the population distributions that describe cell-to-cell variability of single-cell translation kinetics. It was shown that mainly the mRNA half-life is systematically effected by poly(A) length variations. A simple model for the interdependence of mRNA half-life with poly(A) tail length was found enabling further insights into poly(A) tail dependent mRNA degradation. The content of this chapter is based on the manuscript [M3] listed in Associated publications and manuscripts. The study was done in collaboration with Simone Ezendam, who contributed within the framework of her Master's thesis [118], and in cooperation with ethris GmbH.

6.1. Post-transcriptional gene regulation

6.1.1. mRNA decay pathways

The control of mRNA turnover is one of the key parameters for post-transcriptional gene regulation. The degradation of cytosolic mRNA in mammalian cells is mainly regulated by enzymatic removal of adenosine residues at the 3' end, a process called deadenylation [190]. Two major pathways for mRNA degradation were identified that are both initiated by poly(A) tail removal [191]. These pathways are sketched in Figure 6.1.. After shortening the poly(A) tail to a certain length the mRNA is finally degraded following one of two possible pathways. The removal of the poly(A) tail seems to be the rate limiting step for degradation [192]. Firstly, the 5' cap is enzymatically hydrolyzed, resulting in a decapped mRNA. The decapping of the 5' end enables the binding of a 5' exonuclease that degrades the mRNA body in 5' to 3' direction. Secondly, a protein complex called the 3' exosome (containing an exonuclease) binds to the 3' end after deadenylation and degrades the mRNA body in 3' to 5' direction. Less likely pathways independent of deadenylation are the decapping for non deadenylated mRNAs resulting in 5' exonuclease activitiy, or the direct degradation by the 3' exosome.

Beside the mentioned mRNA decay pathways, degradation can also occur through endonucleotic cleavage. In this case, the mRNA body is cleaved somewhere and the thereby generated new 5' and 3' ends are recognized by exonucleases [190]. Endonucleases can be mediated for example by siRNA or a non-sense stop codon within the ORF. As the exoge-

nous mRNAs in this study have an artificial ORF sequence, it is unlikely that edonuclease activitiy is relevant for degradation of these mRNAs.

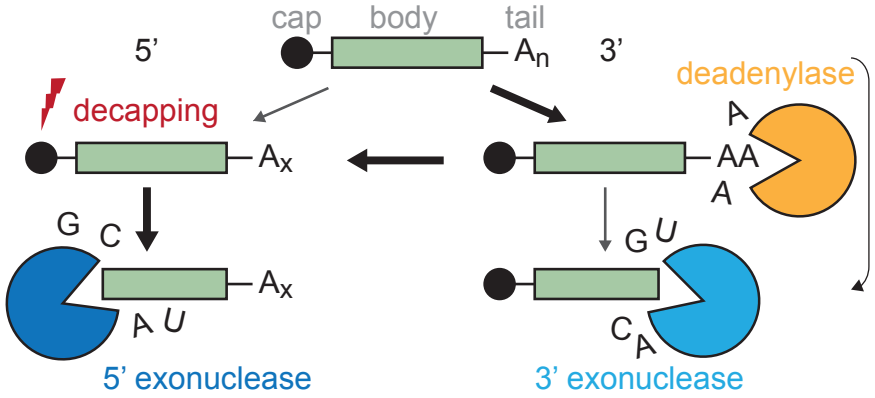

Figure 6.1.: Main mRNA degradation pathway dependent on exonucleases. The major pathway is marked by the thick arrows triggered by poly(A) tail shortening. Less likely pathways or steps are marked by the thin arrows.

6.1.2. Deadenylation regulation

As poly(A) tail shortening is expected to be the rate limiting process for mRNA degradation, a closer look at the key players for deadenylation is given. The naturally occurring key players are RNA sequence elements, deadenylase protein complexes, and poly(A) binding proteins (PABPs) [190]. As the used exogenous mRNAs do not have any UTRs and most RNA elements influencing deadenytlation are encoded in the UTR, RNA elements are not further discussed [38]. Therefore, the regulation can be attributed to the interplay of PABPs and deadenylase complexes. The PABPs have two main functions, whereat multiple PABPs can bind to the poly(A) tail at the same time [193]. The binding of PABP to the tail protects it from degradation and the PABP can furthermore bind to proteins that are bound to the 5' cap themselves. The interaction between cap and tail binding proteins is called circularization and is important for translation regulation e.g. ribosome recruitment and further protects both ends from degradation, therefore enhancing mRNA stability [194, 195]. The deadenylases on the other hand shortens the tail, which finally leads to degradation of the mRNA body by exonucleases. The process of "default" deadenylation, which is independent of any RNA element, consits of two consequent phases with different rates dependent on length [190]. For poly(A) tail lengths larger than 110 nt the protein complex Pan2/Pan3 is active and for smaller poly(A) lengths the protein complex ccr4/caf1 carries out the poly(A) tail shortening until only a few nucleotides are left [190]. During the first phase, a degradation of the mRNA body is not reported. Its degradation is mainly observed during or after the second deadenylation phase [190]. Furthermore, the two deadenylase complexes can interact with each other and form a super complex which enables the coordination of both phases. This super complex can be recruited to the tail as Pan3 and PABP can interact, which in principle can occur at any PABP [190].

Due to this biphasic mechanism of deadenylation, the mRNA half-life is dependent on the poly(A) tail length. The naturally occurring poly(A) lengths for mammalian cells determined by genome wide bulk studies are in the range of 0-300 nt, whereas the median tail length varies for different cell types between 65-95 nt [196, 197]. The abundance of poly(A) tail lengths larger than 250 nt is rare [197]. The asymmetric length distributions for mammalian cells have their maximum around 30-50 nt depending on the cell type. Thereby, the median poly(A) tail length of different genes are around 60 nt with a range of 30-100 nt for 95% of the investigated genes [196].

6.1.3. Exogenous mRNA production of different poly(A) tail lengths

In order to study the impact of different poly(A) tail lengths on protein expression, ten *in vitro* transcribed mRNAs covering the physiological range of poly(A) tail length were synthesized. The mRNA constructs for this study were produced and provided by Zeljka Trepotec from ethris GmbH. The ORF of d2eGFP was cloned in the backbone of a plasmid as described previously [107]. All mRNA constructs of different poly(A) tail lengths are synthesized by IVT using the same protocol. Thereby, the different poly(A) tail length n_A were obtained using two different techniques.

The shorter lengths of n_A =(0 nt, 20 nt, 58 nt, 80 nt, 125 nt) were encoded in the pDNA template after the ORF. Therefore, the synthetic poly(A) sequences of the desired length were introduced into the vector backbone by PCR [198]. During IVT the mononucleotide stretches of length n_A are transcribed by the RNA polymerase generating the poly(A) tail at the 3' end of the produced mRNA. The transcribed mRNAs of the template encoded poly(A) tails are ready to use for transfection studies.

The longer tail lengths of n_A =(60 nt, 125 nt, 190 nt, 240 nt, 290 nt) were synthesized in a separate step after IVT of the mRNA construct without poly(A) tail by post-polyadenylation. Different poly(A) tail lengths were generated by post-polyadenylation and the enzyme poly(A) polymerase was mixed with the construct of n_A =0 nt and adenosines. In order to obtain five different poly(A) tail lengths, the reaction time for post-polyadenylation was varied from 5 min up to 1 h. The poly(A) tail lengths of all constructs were verified by gel electrophoresis and automated capillary electrophoresis.

Both of the two mentioned techniques have their limitations [198]. The encoded poly(A) tails are limited to a length of approximately 120 nt. This is because long homopolymeric stretches like the poly(A) tail tend to recombine during amplification of the plasmid leading to shorter plasmid encoded tails than intended. Furthermore, the error rate of RNA polymerase during IVT increases with encoded poly(A) length as it gets more likely for the a polymerase to slip off the template for long homopolymeric stretches resulting in shorter tail lengths as well [199]. For the enzymatically fabricated poly(A) tails it is possible to generate tail lengths much longer than 120 nt. However, it is difficult to control the final length and the homogeneity of the poly(A) tail length for a certain reaction period [198, 200].

6.2. Translation kinetics depending on poly(A) tail length

6.2.1. Single-cell expression time courses

Each of the exogenous mRNA constructs of the five template encoded poly(A) tail lengths (0 nt-125 nt) and the five enzymatic poly(A) tails (60 nt-290 nt) was analyzed using the single-cell translation assay described in section 2.3. All mRNA constructs were encapsulated with L2000 with a constant ratio of 2.5 µl L2000 per 1 µg mRNA. As the constructs differ in their molecular weights due to the varied poly(A) length (∼24% deviation between the shortest and the longest construct) it has to be considered that the total number of mRNAs, which are accessible for lipoplex formation, varies. However, it is to be expected that this only has an impact on the number of transfected molecules and that the parameter of main interest, the mRNA half-life, is unaffected by the number of transfected molecules. Several experimental replicates were performed using HuH7 cells to measure all constructs at least twice.

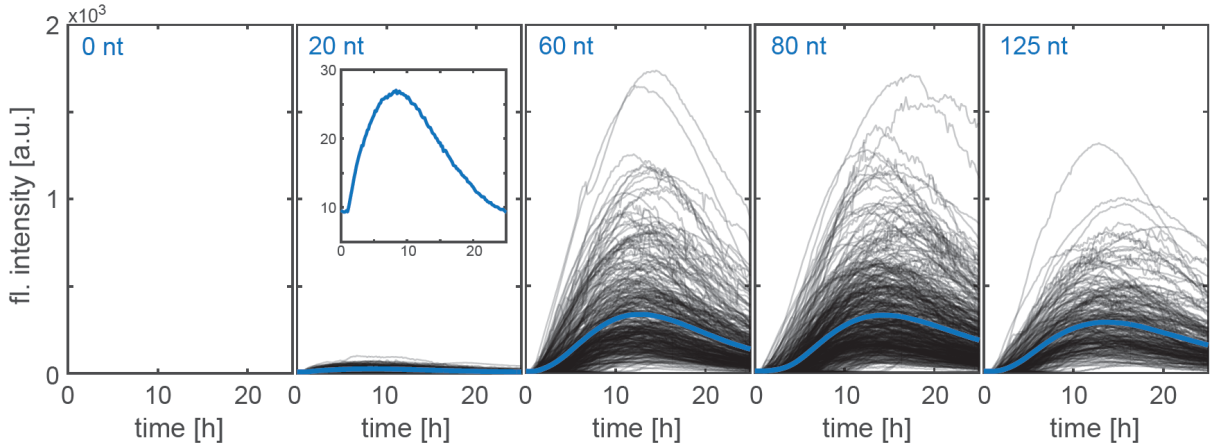

Figure 6.2.: Single-cell time courses of cells transfected with mRNA encoding for d2eGFP with template encoded poly(A) tails. The single-cell time courses (gray) are plotted with the respective mean time course (blue) of different poly(A) tail length (upper left corner: specific length of poly(A) tail).

Representative single-cell time courses of d2eGFP expression of one replicate are shown in Figure 6.2. for the encoded poly(A) tail constructs. The single-cell traces (gray) show a large cell-to-cell variability as observed in the studies previously discussed. Clear differences in the expression strength are visible for different poly(A) tail lengths. No protein expression was detectable for cells transfected with n_A=0 nt. As mRNA is expected to be transfected (all other mRNAs transfected in the same experiment showed protein expression), one can conclude that the mRNA is non functional. The time courses of cells transfected with the short 20 nt tails have low expression levels compared to the time courses of longer poly(A) tails. The 20 nt time courses are very noisy, which is illustrated by the still noisy mean time course (blue trace). However, the mean time course shows a characteristic shape for d2eGFP expressing cells (as observed in Figure 4.1. B as well).

The mean time courses indicate the population behavior and illustrate the different expression levels, which are strongly increasing and less noisy for longer tails compared to the construct with the 20 nt tail.

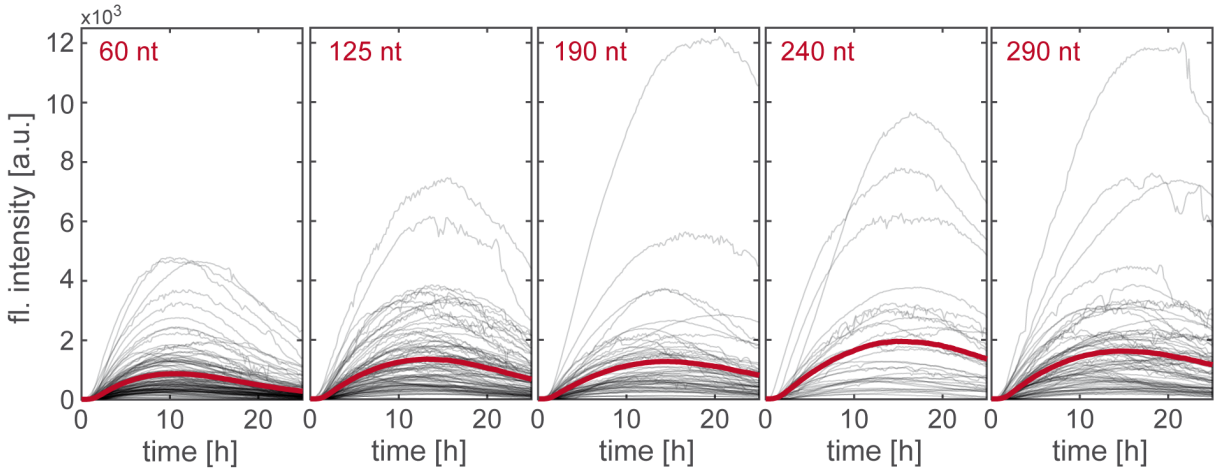

Figure 6.3.: Single-cell time courses of HuH7 expressing d2eGFP after mRNA transfection with enzymatically added poly(A) tails of varied lengths. Each gray time course corresponds to one cell. The population kinetic is shown as mean time course (red line) with the respective poly(A) tail length in the upper left corner.

The single-cell fluorescence time courses for the enzymatic poly(A) tail constructs show a high heterogeneity regarding the translation kinetics in all experimental replicates as representatively shown for one replicate in Figure 6.3.. The expression levels change depending on poly(A) tail length and show a large heterogeneity as expected. Interestingly, the mean expression time course decreases for the construct with 290 nt compared to the construct with the highest mean expression (240 nt length). In general, the fluorescence intensity time courses, which are directly proportional to the protein abundance, increase with longer poly(A) tail lengths, which was reproducible for all replicates. This supports the hypothesis that the mRNA half-life correlates with the poly(A) tail length.

6.2.2. Single-cell parameter estimation

In order to estimate the parameters which describe translation, the analytical solution of the translation-maturation model (Equation 5.2) was fitted to each of the single-cell time courses. Therefore, the same algorithm was used as for the parameter estimation in chapter 5 and the respective algorithm is deposited at zenodo (doi:10.5281/zenodo.2626006). Analog to the fitting of the eGFP time courses in chapter 5, the maturation rate k_M and the protein degradation rate β were fixed to a population value determined in a separate experiment. It was assumed that the maturation rate is the same for eGFP and d2eGFP as the proteins only differ at the N-terminus and the amino acids forming the fluorophor are the same for both GFP variants [99]. For the protein degradation rates it was shown that they differ by a factor of three (chapter 4) and therefore the degradation rate of

d2eGFP was experimentally determined using the mRNA constructs of this study. The destabilized version was chosen because the higher turnover rate of d2eGFP allows for a more direct monitoring of the expression kinetics compared to eGFP [117]. In order to determine the mean protein degradation rate for d2eGFP, a translation block experiment was performed as described in detail in Appendix A. The translation of cells transfected with mRNA encoding for d2eGFP was blocked 8 h after transfection by adding CHX. The fluorescence time courses monitored after translation inhibition (as shown in Figure A.4.) were fitted by an exponential decay to determine the degradation rate. The mean value of the estimated population distribution was used as fixed parameter for β =0.23 h^{-1} for each fit. In Figure 6.4 A representative time-series of d2eGFP expression are shown with the respective fits for the translation-maturation model (blue traces). The model describes the single-cell kinetics in general. However, the residuals reveal that the d2eGFP kinetics show systematic deviations (Figure 6.4 B) between fit and experimental data. Furthermore, the d2eGFP kinetics are not as well described by the model as the eGFP kinetics in chapter 5 (Figure 5.2).

In the previous study described in chapter 4, it was found that the model candidates typically describe the eGFP kinetics better than the d2EGFP (Figure 4.5. D). The ribosome model described the expression kinetics for both reporters reasonable well (equations given in Appendix B). For this reason the ribosome model was taken into account for analyzing the d2eGFP kinetics of varied poly(A) lengths. In order to keep the number of free parameters as low as possible, the ribosome binding kinetic was approximated as Michaelis-Menten kinetic [118]. This approximation leads to a simplified ribosome model referred to as Michaelis-Menten model:

$$\begin{aligned} \frac{d[mRNA]}{dt} &= -\delta[mRNA],\ [mRNA](t_0) = m_0 \\ \frac{d[GFP]}{dt} &= -\beta[GFP] + k_{TL}\frac{[mRNA]}{1+\alpha[mRNA]},\ [GFP](t_0) = 0 \end{aligned} \tag{6.1}$$

The translation rate k_{TL} is given by the quotient of the reaction's maximal speed v_{max} with the Michaelis-Menten constant k_{MM}. The parameter $\alpha = {}^1/_{k_{MM}}$ describes the ribosome activity. The difference between the Michaelis-Menten model and the two-stage model (Equation 4.1) is that it incorporates the ribosome kinetics into the translation step but is otherwise the same. However, no analytical solution could be determined, which is why the solution for the parameter estimation had to be solved numerically for each time course. In Figure 6.4. A the respective fits of the Michaelis-Menten model (green) are plotted together with the experimental time courses (gray) and the fits for the translation-maturation model (blue). The Michaelis-Menten model described the experimental data slightly better which was verified by the lower residuals of the Michaelis-Menten model (Figure 6.4. C). However, the Michaelis-Menten model did not neither resolve the systematic deviations of the fits to the experimental data. Furthermore, the Michaelis-Menten model has two more free parameters than the translation-maturation model and has to be solved numerically, which is computationally more expensive. For these reasons, the

parameter estimation was performed using the translation-maturation model taking systematic deviations into account.

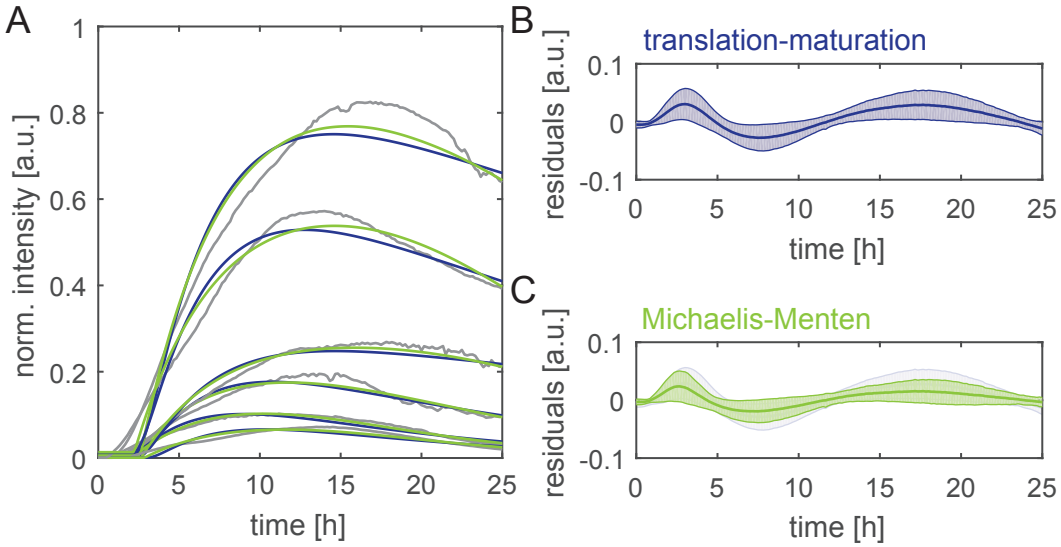

Figure 6.4.: Single-cell expression time courses of d2eGFP. (A) Representative fluorescence time courses with the corresponding fits of the translation-maturation model (blue) and the Michaelis-Menten model (green). (B) Residuals for both models are shown as mean ±std. The residuals of the translation-maturation model are shown in the background of the Michealis-Menten model in lighter color.

6.3. Influence of poly(A) tail length on single-cell translation kinetics

6.3.1. poly(A) tail independent parameter distributions

Four parameters are determined by fitting each cell: the fluorescence offset, the onset time t_0, the expression rate $m_0 k_{TL}$, and the mRNA degradation rate δ. The offset and the onset time are expected to be independent of the tail length variation. The offset parameter has biologically no meaning and mainly depends on image acquisition settings e.g. the exposure time. The distributions of the offset parameter did not show systematic changes for mRNAs of different length. This is shown in Figure 6.5. A for one representative experiment of cells transfected with the enzymatically fabricated poly(A) tails. All offset distributions of different replicates or the template encoded mRNAs did solely show random variations. The protein expression onset depends on the mRNA delivery system as well as the environmental properties like FBS concentration during transfection (see chapter 5). All mRNA constructs were complexed using the same transfection protocol for L2000. All mRNA lipoplex samples were incubated without FBS for one hour. For these reasons, it is expected that the onset time distributions for mRNAs with different poly(A)

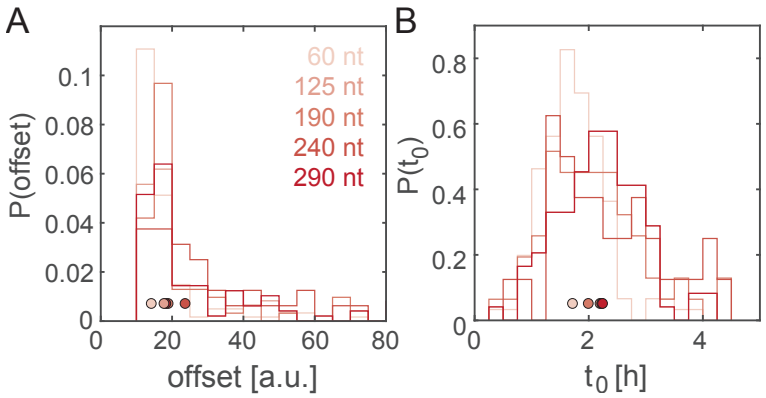

Figure 6.5.: Distributions of offset (A) and onset time (B) for mRNAs of different poly(A) tail lengths with color intensity indicating the tail length. The median value for each distributions is shown as a dot in the respective color.

tail length do not show substantial deviations. This expectation was confirmed for the onset time distributions of all replicates and is exemplarily shown for one experiment in Figure 6.5. B.

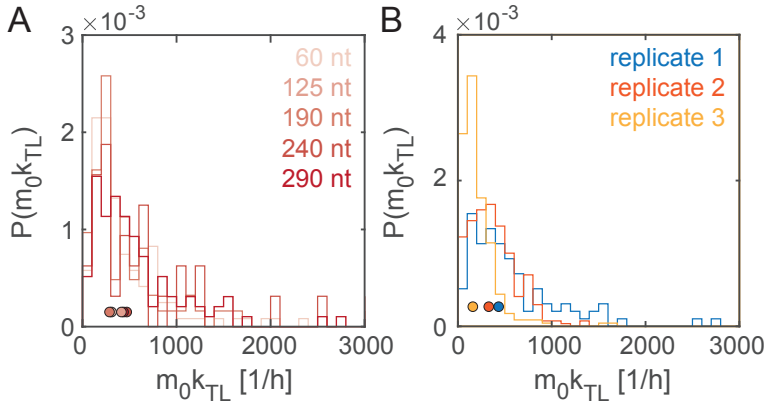

Figure 6.6.: The distributions of the expression rate $m_0 k_{TL}$ show the same variations depending on poly(A) tail length (A) as for replicates of 125 nt length (B). The median value of each data set is represented by a dot with color intensity indicating the tail length or rather the replicate.

Beside the two discussed parameters, the expression rate $m_0 k_{TL}$ did not show a systematic shift for increasing poly(A) tail length as well (Figure 6.6. A). The median expression rates of one experiment for the five different enzymatically synthesized poly(A) tail lengths vary between 292 h^{-1} and 474 h^{-1} in a non-systematic matter. The variation within the experiment could arise from different amounts of transfected mRNA molecules m_0 as each of the L2000-lipoplexes had to be prepared separately. The random deviations within a replicate are in the same range as for cells treated with the same mRNA construct at different days (Figure 6.6. B). The observed variation for the expression rate is likely due to batch effects that possibly arise from variations within the transfection process like

discussed in section 4.4 (see Figure 4.6. D).

6.3.2. mRNA half-life changes with increasing poly(A) tail length

The final parameter, which is expected to be dependent on poly(A) length, is the mRNA degradation rate. For the following results the degradation rate is converted into mRNA half-life $\tau_{mRNA} = {}^{ln(2)}/_{\delta}$. The single-cell data of the results shown so far were not pooled to discuss the intra- and interexperimental variance of the data. In order to increase statistics, the data was pooled for each mRNA construct after checking whether the results were reproducible.

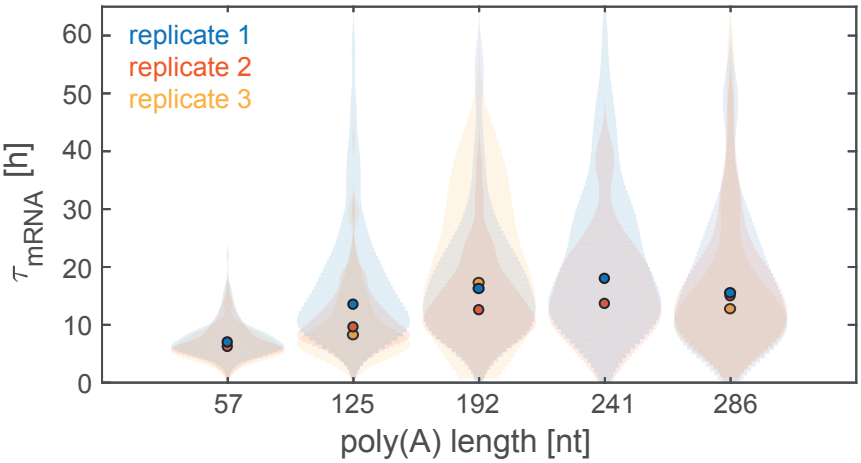

Figure 6.7.: Day-to-day variance of the half-life distribution estimated as Kernel densities with the respective median values represented as dots depending on poly(A) tail length. The coloring corresponds to different replicates.

Figure 6.7. shows the overlay of the mRNA half-life distributions for three replicates of the enzymatic poly(A) constructs. Most distributions of the same mRNA construct cover a similar range across replicates indicating that pooling of the data is appropriate. Some individual distributions (as the distribution of replicate 1 at of 125 nt) show a shift to higher half-lives as the other two replicates. However, the trend of the median half-lives is reproducible for all replicates. The same was observed for the other four template encoded constructs. For this reason, the data sets of all replicates were pooled for further analysis.

Figure 6.8. shows the single-cell data for each construct with the estimated population distribution estimated in gray. The respective median with the first and third quartile are plotted for clarification. No mRNA half-life was determined for n_A=0 nt as no d2eGFP expression was measured. Of course, the mRNAs without a poly(A) tail have a physical half-life but as they lead to no expression the functional half-life is zero. For the four constructs of template encoded poly(A) tails (left plot of Figure 6.8.) the half-life distributions were shifted to higher half-lives with increasing tail lengths of almost a factor four between the shortest tail of 20 nt and the longest one with 125 nt. Beside the shift to higher half-lives, the estimated Kernel distributions broadened from a FWHM of 3.6 h

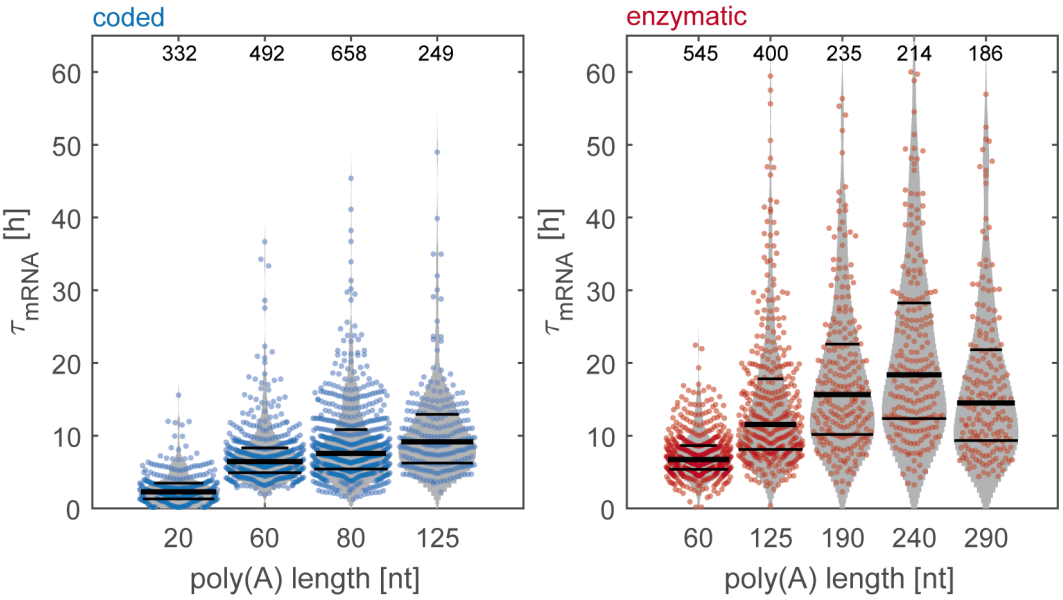

Figure 6.8.: The dependency of poly(A) tail size with mRNA half-life at the single-cell level of the pooled data sets. The left plot shows the pooled data of all experiments with the four constructs of encoded poly(A) tails. The five enzymatically added poly(A) tail constructs are shown in the right plot. Each data point correspond to a single cell. The black lines mark the median value (thick line) and the interquartile range of each construct. The gray areas correspond to the estimated Kernel distributions.

to 9.9 h. For the enzymatically added poly(A) tails (right plot of Figure 6.8.) the distributions shift to higher half-lives as well except for the construct with the longest tail of 290 nt. The decrease may be biologically relevant as it is reported that hyperpolyadenylated mRNAs can result in faster decay as longer is not always better [38, 201]. The maximal increase for the enzymatic poly(A) tails is about a factor of three between the tail with n_A=60 nt and n_A=240 nt. The mean half-life for the longest tail decreased to the same level than the mRNA construct of n_A=190 nt. The trend of the distribution widths (FWHM) is the same than for the mean values. The FWHM increases up to 22.3 h for n_A=240 nt. The mean values and FWHM for all constructs are given in Table 6.1..

The poly(A) lengths were designed to overlap each other for the template encoded and enzymatically added tails. The poly(A) lengths of 60 nt and 125 nt were analyzed for both fabrication methods. In Figure 6.9. the half-life distributions for the constructs of same theoretical lengths are shown. The half-life distributions of template encoded tails (blue) and the enzymatically added tails (red) are placed opposite to each other. For the shorter length of 60 nt the distributions are almost identical with a divergence of the mean values smaller than the standard error of the mean (see Table 6.1.). This is not the case for the distributions of 125 nt length. The half-life distribution of enzymatic tails is shifted to higher values with a divergence of 4 h for the mean values while the standard errors are in a similar range of 0.4 h and 0.5 h respectively. This occurs most likely due to the different fabrication methods. It is expected that the enzymatically generated poly(A) tails lead to broader tail length distribution with increasing reaction time as it is an enzymatic

Table 6.1.: Mean half-life of mRNA constructs of template encoded (c) and enzymatically added (e) poly(A) tails with the respective standard error of the mean.

length [nt]	mean half-life [h]	sem [h]	FWHM [h]
20 c	2.73	0.12	3.62
60 c	7.35	0.19	5.46
60 e	7.31	0.13	4.98
80 c	8.96	0.22	6.96
125 c	10.48	0.40	9.93
125 e	14.48	0.48	10.90
190 e	17.91	0.68	17.55
240 e	21.32	0.83	22.28
290 e	17.32	0.81	17.08

reaction compared to template encoded poly(A) tails [200]. Furthermore, the template encoded length of 125 nt is possibly shorter in reality due to spontaneous deletion of parts of the encoded homonucleotide stretches during plasmid amplification in bacteria [198]. These reasons could explain the differences for the half-life distributions as it is expected that the poly(A) tail length distributions for the two fabrication methods are not the same for longer poly(A) tails.

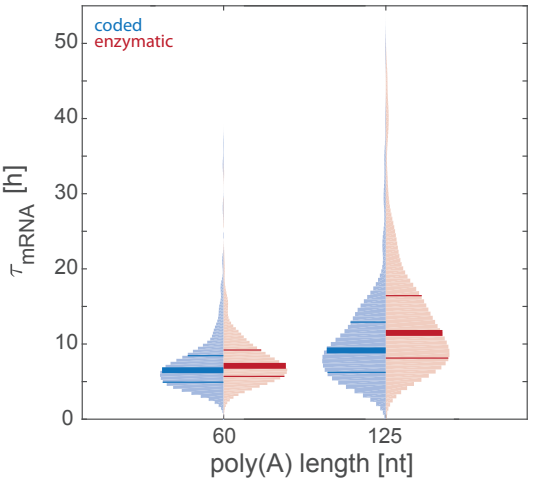

Figure 6.9.: Half-life distributions of mRNA constructs with template encoded (blue) and enzymatically added (red) poly(A) tails. The estimated Kernel densities are plotted with the median (thick line) and the interquartile range (0.25 quantile and 0.75 quantile).

Figure 6.10. A shows the mean half-life of each mRNA construct as a function of poly(A) tail length with the respective standard error of the mean. The template encoded mean half-lives (blue) increases depending on longer poly(A) tail length. The enzymatically fabricated poly(A) tails (red) show increasing half-lives as well. However, the half-life of the construct with the longest tail (290 nt) has a lower half-life than the construct with n_A= 240 nt. In order to verify if the decreased half-life has biological origin further

investigations are needed. The mean half-lives follow the same trend as discussed for the population distributions (Figure 6.8.).

Next, a simple model was developed that describes the mRNA half-life τ_{mRNA} as a function of the poly(A) tail length n_A. It was assumed that a critical poly(A) tail length n^* is needed for translation initiation in order to enable eGFP expression. Furthermore, it is assumed that the mRNA decay is sequentially depending on the poly(A) tail length and that this dependency is non-linear. As a first ansatz, the dependency of the mRNA half-life τ_{mRNA} on the poly(A) tail length n_A is described by a power function depending on a pre-factor α and an exponent ε:

$$\tau_{mRNA} = \alpha \cdot (n_A - n^*)^{\varepsilon} \tag{6.2}$$

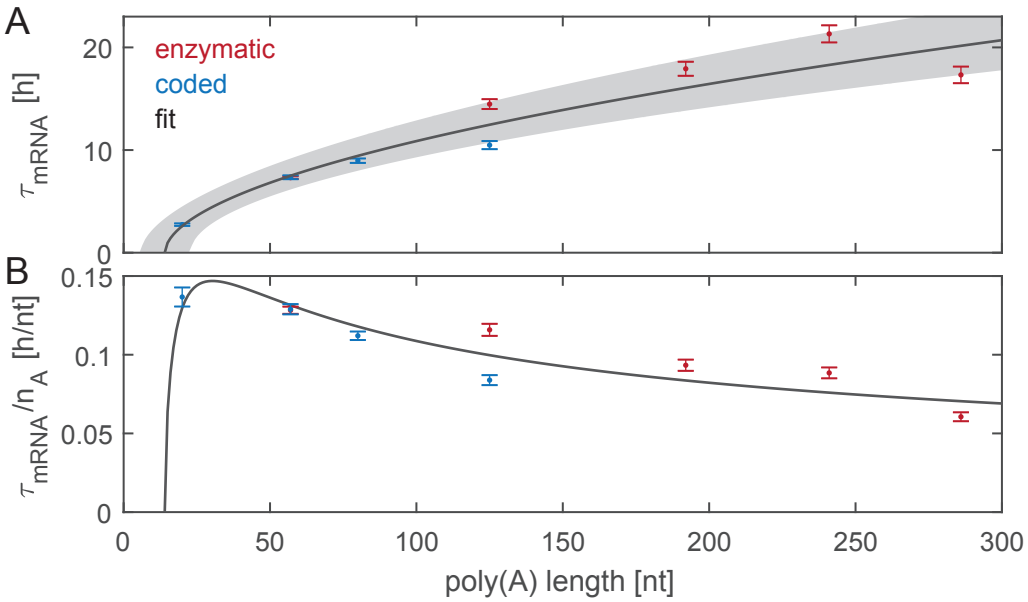

Figure 6.10.: Influence of the poly(A) tail length on the mean mRNA half-life τ_{mRNA}. (A) The mean half-life for each construct with template encoded (blue) and enzymatically added (red) poly(A) tails are plotted with the respective sem. The gray line corresponds to a fit with the 95% boundaries. (B) The normalized mRNA half-life to the poly(A) tail length n_A with the normalized experimental data.

The function given by Equation 6.2 was fitted to the data shown in Figure 6.10. A. The mean values were weighted based on the standard error to fit the model (Equation 6.2) to the data. The fit (gray trace) illustrates that the model is able to describe the experimental observations. The dashed lines mark the 95% confidence boundaries of the fit for an estimation of the error. The following fit parameters were determined: α=1.1, n^*=14.1, ε=0.53. In order to investigate at which poly(A) tail length the impact of a single adenosine is the highest, the mRNA half-life was normalized by the poly(A) tail length. The normalized mRNA half-life was plotted as function of the poly(A) tail length together

with the experimental data (Figure 6.10. B). The normalized mRNA half-life shows a fast increase for poly(A) tails longer than the critical length n^* which peaks around a length of 30 nt and decreases much slower afterwards. The length at the half maximum of the normalized half-life smaller than 30 nt is reached at a length of 16 nt. In contrast, the length of half maximum larger than the maximum is reached at 260 nt illustrating the asymmetric trend. The maximum n_{max} of 30 nt is dependent on the critical length n^* and the exponent ε :

$$n_{max} = -\frac{n^*}{\varepsilon - 1} \tag{6.3}$$

Interestingly, a poly(A) tail length of 30 nt corresponds to the most frequent poly(A) tail length found in several human cell lines [197].

6.4. Discussion

LISCA experiments of cells transfected with ten different mRNA constructs of varied poly(A) tail length enabled single-cell protein expression and the accurate measurement of functional mRNA half-life distributions. A high heterogeneity of expression levels was observed for each mRNA construct that yields to reporter protein translation (Figure 6.2. and Figure 6.3.). This finding already indicates that at least one of the parameters that describe the translation kinetics is proportional to the poly(A) tail length. The parameters of each cell were determined using the established translation-maturation model. The Michealis-Menten model lead to slightly better description of the experimental time courses but at the cost of two additional fitting parameters (Figure 6.4.).

The parameter distributions for all experiments were estimated and it was investigated whether they change systematically for increasing poly(A) tail length. The offset parameter and the onset time appeared to be unaffected, as expected (Figure 6.5.). The expression rate did not show a dependency on poly(A) tail length neither (Figure 6.6.) which could have been possible as PABPs interact and stabilize proteins bound at the 5' cap and it is known that this leads to enhanced translation initiation [186]. It is conceivable that the expression rate $m_0 k_{TL}$ shows correlation with poly(A) tail length. In literature it is reported that translation efficiency sometimes is coupled with poly(A) tail length. Subtelny et al. showed that this is the case for translation in embryonic cells, but that the effect dissolves during development and was not found for differentiated cells [197]. In a different study of Chang et al., it was observed that the translation efficiency in human cell lines did not correlate with poly(A) tail length, which is in agreement with Subtelny et al. as they used cell lines e.g. HeLa, which are differentiated cells [196]. As the used HuH7 cells are non-embryonic liver cells, it is possible that the translation rate is unaffected by increasing poly(A) tail length as well. However, as the translation rate cannot be estimated as single parameter it is possible that the large variation of transfected mRNAs m_0 masks changes within the translation rate k_{TL}.

It was shown, that the parameter that changes as a function of poly(A) tail length is the mRNA stability measured by the functional mRNA half-life (Figure 6.7. and Figure 6.8.). The correlation of the mean half-life with poly(A) tail length in Figure 6.10. A shows a clearly non-linear trend. The dependency is described systematically for the first time. A power function model (Equation 6.2) describes the experimental data reasonably well. The results of the fit give new insights in the regulation of poly(A) dependent mRNA degradation. A critical length of about 14 nt is derived from the experimental data that is needed to observe eGFP expression. Interestingly, this is the same length scale that is reported to bind one PABP [202]. However, at least two PABPs covering a length of 27-30 nt are needed for stable translation initiation [203]. This is one explanation why the shortest length of 20 nt has a low expression level resulting in a low functional stability. In addition, it is reported that poly(A) lengths shorter than ∼25 nt are immediately enzymatically marked which functions as 'final' signal for mRNA degradation [204]. The length for binding two PABPs is exactly the length at which the normalized mRNA half-life has its maximum and prevents the mRNA from marking for the final decay (Figure 6.10. B). The impact for one nucleotide in the poly(A) tail range between the critical length and 30 nt increases very fast which could be due to an increasing probability for a second PABP to bind with increasing length up to 30 nt. Therefore, the single-cell studies confirm that the binding of two PABPs is an important benchmark for mRNA stability. The poly(A) dependent mRNA degradation process has to be considered as a dynamic interplay between cap binding proteins, the PABP and the protein complexes containing deadenylases (Figure 6.11.).

The deadenylation process occurs in two length dependent phases [190]. It is reported that the deadenylation happens at different rates depending on the active deadenylase complex (as described in subsection 6.1.2.). For tails longer than 110 nt the Pan2/Pan3 complex is active until a length of 70-80 nt after the protein complex of the second phase takes over. The complex leads to a rather slow poly(A) tail shortening as it is reported that the Pan2/Pan3 complex shows a distributive digestion of the poly(A) tail [190]. A distributive digestion is characterized by removing only a few nucleotides before the enzyme slips off the target and has to bind to a new one [205]. The second phase, in contrast, is reported to show a processive digestion kinetics. During the second phase the complex ccr4/caf1 is actively binding to poly(A) tail length in the range of 10-110 nt. For the processive digestion by ccr4/caf1, the complex remains bound to the target and removes nearly the whole poly(A) tail before releasing the target [205]. This would result in an linear dependency between mRNA half-life and poly(A) tail length up to 70 nt if only the processive degradation takes place. However, this was not resolved in the experimental data. Furthermore, there is experimental evidence that for default deadenylation the complexes of both phases form a "super" complex that enables the coordination of both phases [206]. The super complex possibly gets recruited to the poly(A)-PABP complex as Pan3 is able to interact with PABP [190]. The biphasic behavior of deadenylation explains why the mRNA half-life does not correlate with poly(A) tail length in a linear way but depends on slow distributive digestion for at least two thirds of the investigated poly(A) tail range. Distributive enzyme kinetics are limited by the diffusion of the enzyme to the target and the rates typically correlate with the abundance of target n by $n^{0.5}$. A

similar exponent of ε=0.53 was found for the model describing the dependency of poly(A) tail length on mRNA half-life, which is an indication that the stability regulation of the investigated constructs is dominated by a distributive poly(A) digestion.

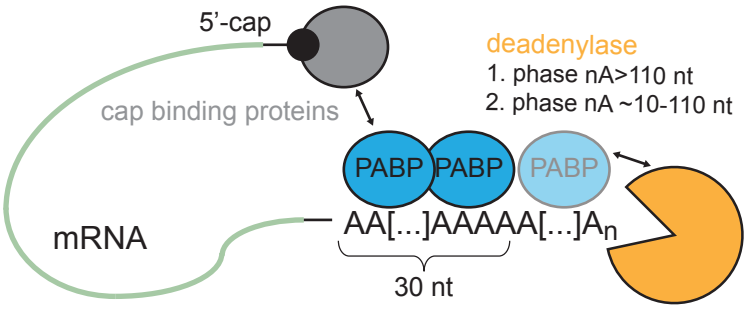

Figure 6.11.: Schematic mRNA deadenylation depending on poly(A) tail length. The dynamic process of deadenylation depends on the interplay between the cap binding proteins, the PABP, and the deadenylases shortening the poyl(A) tail. Two different deadenylase complexes are active in a tail length dependent manner.

In conclusion, a first systematic description how poly(A) tail length affects mRNA half-life was given by the proposed model that described the single-cell data (Equation 6.2). This finding is a further verification that single-cell translation studies enable new insights into protein expression kinetics. The approach is accurate enough to measure the cell-to-cell variability within a population and provides enough statistics to determine mean values with a small error rate which enabled further investigation of the mRNA half-life as a function of poly(A) tail length.

7. Dual-color approach for combinatory dynamic analysis

Single-cell protein expression time courses proved to be a suitable readout to quantify translation after mRNA transfection to give new insights e.g. in mRNA nanocarrier transport as shown in the previous chapters 4-6. In order to gain more information from LISCA studies, the readout of two or more fluorescent signals within one cell had shown to be beneficial. Other studies investigated the time dependence of apoptotic events at the single-cell level using fluorescent marker combinations to determine the occurrence of intermediate steps of programmed cell death [75, 184]. Depending on the marker combination, this enabled the correlation of event times or whole dynamics. In this chapter, three approaches are listed describing how the observation of two parallel fluorescence intensity time courses can be applied for single-cell translation studies. Firstly, the expression of two reporter proteins within the same cell can be easily achieved. This is done by the simultaneous delivery of the genetic information for both reporter proteins. Secondly, as the delivery process of a mRNA nanocarrier itself depends on several intermediate steps, fluorescent markers identifying one of the intermediate steps of the transport process would allow for further correlation studies with the potential to reveal the crucial steps of particle uptake. Finally, the direct measuring of the mRNA concentration would be beneficial for the quantification of translation kinetics to improve parameter estimation.

7.1. Co-expression of reporter proteins

Previous studies related to the quantification of gene delivery and expression dynamics observed fluorescence intensity time courses of cells expressing two different reporter proteins [67]. In order to observe two reporter protein expression kinetics within the same cell, the genetic information for both proteins needs to be delivered. Therefore, plasmid DNA (pDNA) encoding for a fluorescent protein or mRNA can be used. If two pDNAs encoding for varied fluorescent proteins of different spectral properties are encapsulated in a nanocarrier, it is possible to observe the fluorescence intensity time courses for both proteins simultaneously. Schwake et al. developed a two-step stochastic model that described the transfection and protein expression process in single-cells [67]. They showed that the model was able to predict the number of transfected plasmids and the distributions of expression level. By investigating co-transfected cells, it could be proved that even the co-transfection rate for the plasmids was predictable by the model. Additionaly, the simultaneous record of two protein expression kinetics can help to break parameter symmetries of the translation model used for fitting as mentioned in chapter 4.

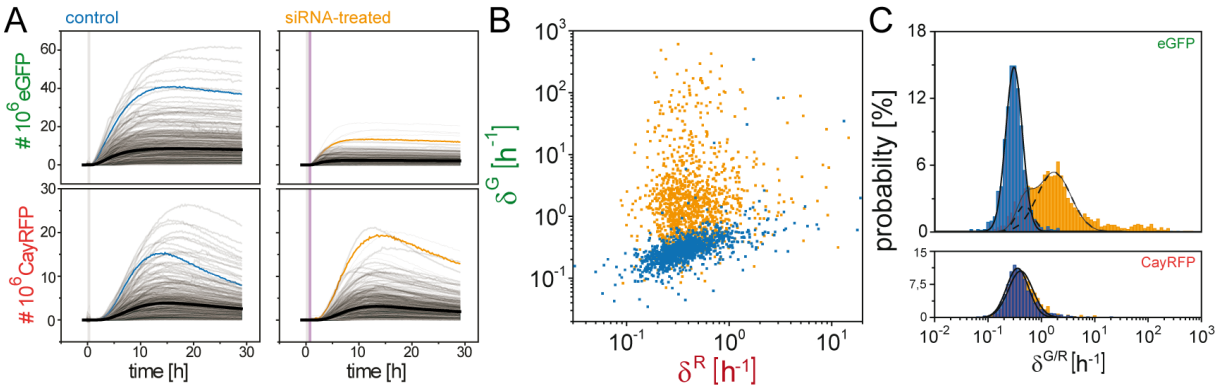

Figure 7.1.: siRNA-mediated silencing of single-cell kinetics. (A) HuH7 cells transfected with two mRNA constructs encoding for eGFP and CayRFP (gray bar) followed by siRNA transfection targeting the eGFP mRNA (purple bar). Cells transfected only with mRNA (left column) as a control and cells transfected with siRNA (right column). Each time course corresponds to a single-cell and the respective mean time course is presented in blue for eGFP or orange for CayRFP. (B) Scatter plot of single-cell degradation rates for eGFP δ^G and CayRFP δ^R of control cells (blue) and siRNA treated ones (orange). (C) The histograms for the degradation rates clarify the distribution shift of siRNA treated cells to higher rates for eGFP revealing a subpopulation structure. The distributions for CayRFP are unaffected. Figure adapted with permission from [119].

Furthermore, the dual expression of reporter proteins can be used to compare differences of expression kinetics in the same cell. This was done for single-cell protein silencing studies, whereas the non-silenced protein expression is used as control for the silenced protein expression [119]. This approach is described in detail in [M4] (listed in Associated publications and manuscripts). In this study, siRNA with a complementary sequence to the mRNA encoding for eGFP was used to study the silencing effect mediated by siRNA binding at the single-cell level. After siRNA transfection, the lead strand of the double stranded siRNA is bound in the so called RNA-induced-silencing complex (RISC). The RISC containing the leading strand is able to bind to a complementary mRNA sequence which induces mRNA cleavage. The cleaved mRNA is recognized by the cell and further degraded and consequently the targeted protein expression is silenced [128]. Many clinical trials investigated the efficiency of siRNAs to downregulate or even knockdown disease related protein expression [12, 207]. The basic principle for dual protein expression is analog to the co-transfection with pDNA [67].

Two mRNAs in equal amounts encoding for two fluorescent proteins were encapsulated in L2000 lipoplexes. The single-cell translation assay is performed as previously described by transfecting the cells with the two different mRNA sequences encoding for eGFP and the red fluorescent protein, so-called CayenneRFP (CayRFP). As the mRNAs are transfected in one step, it is very likely to observe a high percentage of cells expressing both reporter proteins. The kinetics of protein expression differed from each other mainly because the protein degradation rates are not the same. Representative protein expression time courses are shown for both reporters in Figure 7.1. A. Nanoparticles containing both

mRNAs were incubated for 1 h (gray bar) at the beginning of the experiment. A sub set of the cells were transfected with siRNA (purple bar) after mRNA transfection with siRNA specific for the eGFP encoding mRNA. The CayRFP expression time courses are comparable between siRNA treated and untreated cells, the eGFP expression is lower for siRNA treated cells than for the control cells (Figure 7.1. A). The single-cell time courses were analyzed as described in subsection 5.3.1.. As the transfected siRNA is specific for the eGFP coding sequence, it is expected that the mRNA degradation rate distribution changes for the eGFP encoding mRNA depend on siRNA treatment and that the distribution for CayRFP is unaffected. The scatter plot of the mRNA degradation rates determined per cell is shown in Figure 7.1. B. The mRNA degradation rates for CayRFP δ^R scatter over the same range for siRNA treated cells (orange) and untreated cells (blue). This is not the case for eGFP encoding mRNA degradation rates δ^G. The degradation rates for the untreated cells are less scattered than the siRNA treated cells. The histograms for the degradation rates in Figure 7.1. C further illustrate the broader distribution of δ^G after siRNA transfection compared to the distribution of the control cells. The reference of CayRFP shows almost identical distributions independent of siRNA transfection. In order to determine the distributions, log-normal distribution functions were fitted to the histograms. The siRNA treated distribution of δ^G was fitted by two log-normal distribution functions, which describe the data reasonably well.

The most prominent change in the protein expression kinetics after siRNA transfection was the change of δ^G distributions. As the δ^G distribution of siRNA treated cells can be estimated by fitting two log-normal distributions it is likely that the population consists of two subpopulation of cells: One with lower degradation rate, overlapping with the distribution of control cells, and a second subpopulation of higher degradation rates. The low-level population can be partly explained by low siRNA activity (escape of siRNA transfection or low number of delivered siRNA molecules). The most likely explanation of the bimodal distribution is the combination of probabilistic siRNA transfection and its independence of siRNA-unrelated mRNA degradation. Importantly, the shift of the δ^G distributions allows for quantification of the gene silencing efficiency by relative increase in eGFP-mRNA degradation constants as fold change of the median degradation constants. The expression of two reporter proteins proved to be useful as an internal control to investigate siRNA-silencing activity, as the internal control of a second reporter enables to quantifye the siRNA activity and the specificity of the siRNA for the targeted mRNA.

7.2. Endocytosis marker

The use of two fluorescent proteins is a suitable approach for internal controls, but it does not provide further insight in the mRNA delivery process. As the mRNA transport is one of the key challenges for translation studies after mRNA transfection, it is reasonable to have a closer look on the uptake pathway itself. The mRNA transport into the cytosol is carried out by endocytosis (as described in section 5.1.). Thereby, the endosomal escape of the mRNA nanocarriers into the cytosol has been identified as the main bottleneck [135]. The endosomal escape of lipid-based nanoparticles containing mRNA is most likely

due to lipid-fusion of the nanoparticle with the endosomal plasma membrane followed by endosomal leakage or rupture [161]. A variety of assays are reported to investigate the endosomal escape of nanoparticles [161]. But most of them rely on population averaged measurements, usage of artificial endosomes as a model to study the mechanism itself, or rely on high resolution microscopy of single-cells [208, 209, 210]. All of these methods are either not suitable to be combined with the single-cell translation assay or it needs to be verified if an adaption is possible. For this reason two fluorescent markers, which accumulate in endosomes, were tested whether their signal time courses give additional information on the transport process.

Dextran for endosome labeling

Dextrans, which are high molecular weight, branched, and neutral polysaccharides, can be fluorescently labeled and are often used in literature to make endosomes visible due to their endocytotic uptake [208, 209]. Fluorescent dextrans are commercially available in a variety of different spectral properties and molecular weights, one of them is pHodro™ (Thermo Fisher). The presented data is obtained using pHodro Red with a molecular weight of 10,000 Da. The fluorescence emission of pHodro is pH dependent in the range of pH 4-8, which is the relevant range for endocytosis maturation, with high emission under acidic conditions. As the fluorescence intensity is pH dependent, it is possible that the time course of the signal shows a decrease that is related to endosomal leakage or rupture due to nanoparticle escape. The dextran was applied using the manufacturer's protocol [211] and added with the transfection complexes to the cells at a concentration of 25 µg/ml.

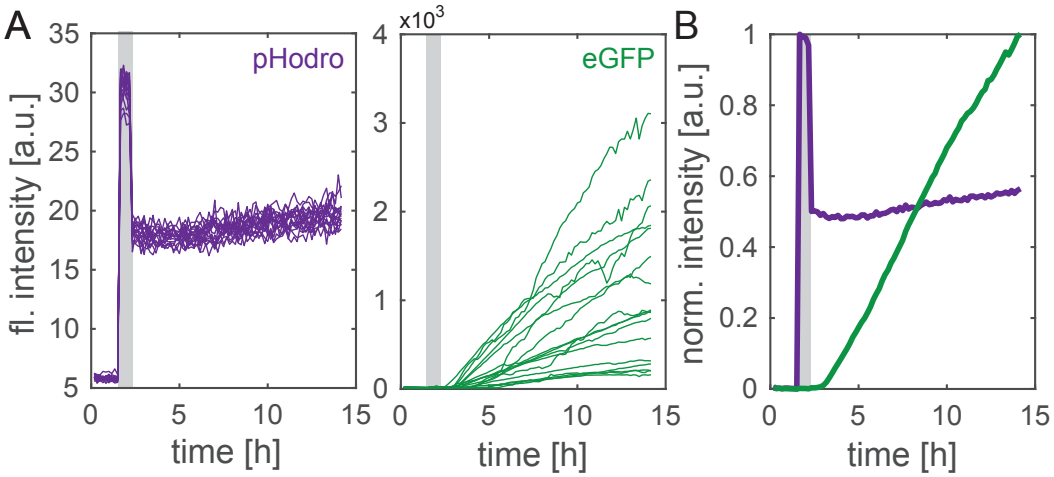

Figure 7.2.: Representative single-cell time courses of cells expressing eGFP after transfection co-labeled with fluorescent dextran. The gray bars mark the mRNA-nanocarrier incubation period. (A) Fluorescence intensity time courses of the endocytosis marker pHodro (purple) and eGFP expression (green) after mRNA transfection. (B) The normalized mean time courses for both fluorophores are shown.

The experiment using pHodro was performed analogously to the previously described

single-cell translation assay by transfecting the cells with mRNA encoding for eGFP during a time-lapse measurement. The single-cell time courses in Figure 7.2. A (left plot) show the fluorescence signal of pHodro. The fluorescence intensity has a sharp increase at the time of dextran addition (the gray bar marks the transfection period). The signal during incubation remained on a maximal level, which possibly overlaid an endocytotic particle uptake or rupture kinetic. After the washing step at the end of the incubation period, the intensity dropped to a certain level and stayed almost constant for the remaining measurement. The single-cell time courses of expressed eGFP showed the expected kinetics and appeared to be unaffected by the additional fluorescence marker as shown in Figure 7.2. A (right plot). The normalized mean fluorescence intensities of pHodro and eGFP in Figure 7.2. B illustrate that the pHodro signal dynamic marked the incubation period of the dextran. The pHodro signal gave no additional information on the particle uptake, which could have been correlated to the eGFP expression kinetics. Therefore, pHodro appeared to be unsuitable as a fluorescent marker to identify intermediate steps of mRNA transport or to obtain any information about the endocytotic activity under the applied conditions.

LysoTracker

As the endosome staining with dextran gave no information on mRNA nanoparticle related endocytotic kinetics, a second fluorescent marker, called LysoTracker, was tested to label acid cell organelles like late endosomes and lysosomes. The LysoTracker (Thermo Fisher) is pH dependent as pHodro and accumulates in acidic organelles [212]. Murschhauser et al. measured LysoTracker time courses in LISCA studies to determine the event time of lysosome permeabilization [184]. The time courses in this study showed an initial increase followed by a slower decrease. During the decrease the signal completely dropped to the background level within a short time frame of 10-20 min. The signal's drop marked the time point of lysosomal rupture because the marker was non-fluorescent at cytosolic pH. This study verified that LysoTracker is a suitable fluorescent marker to record time courses that give information on endosomal/lysosomal activity.

In order to investigate if the LysoTracker is also suitable to resolve mRNA nanoparticle uptake or release, the marker was co-incubated at a concentration of 75 nM during mRNA transfection. Figure 7.3. A (left plot) shows representative single-cell time courses for LysoTracker treated cells. The signals showed an initial increase, longer than the transfection incubation indicated by the gray bar, followed by a decrease, which is slower than the increase period. In contrast to the traces recorded by Murschhauser et al., no signal drop was observed [184]. A signal drop to the background level would indicate an early event of programmed cell death namely lysosomal membrane permeabilization, which is unintended for the transfection studies, and therefore a good indication that the cells are not stressed. However, an early signal drop, which does not reach background level, could have given an indication for endosomal rupture of some organelles due to mRNA escape from the endosomes. The eGFP expression time courses in Figure 7.3. A (right plot) showed the expected kinetics that are comparable to the ones of pHodro treated cells (Figure 7.2. A), which indicates that the mRNA translation is not influenced

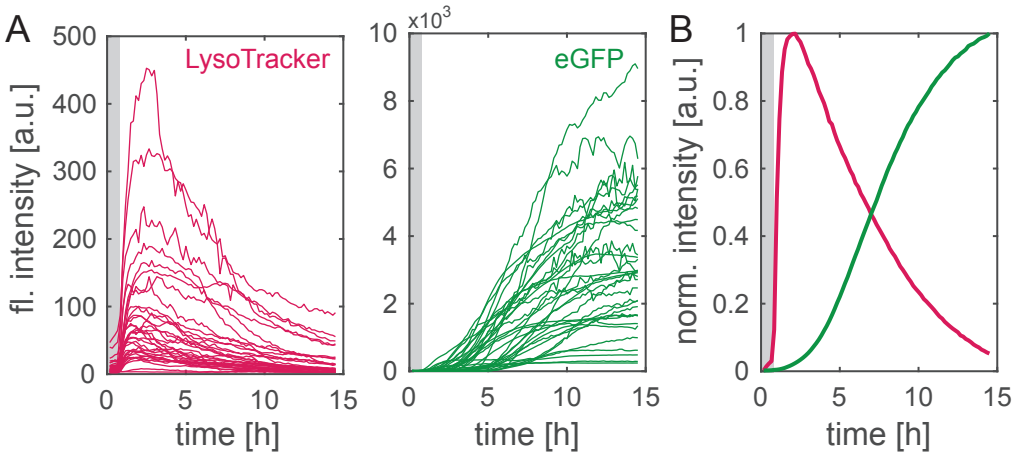

Figure 7.3.: HuH7 cells transfected with eGFP encoding mRNA. (A) Single-cell kinetics of the endosome marker LysoTracker (pink) and the protein expression of eGFP (green). The gray bar marks the mRNA-nanocarrier incubation period. (B) The normalized mean time courses for Lysotracker (pink) and eGFP (green) with fluorescence intensities normalized to the respective maximum of the time course.

by LysoTracker. The normalized mean time courses for LysoTracker and eGFP illustrate the population behavior (Figure 7.3. B).

Conclusion

The two investigated endosome markers, pHodro and LysoTracker, did not show any features in their respective time courses that appeared to be suitable to extract an event time for endosome uptake or rupture or any other information under the tested conditions. As endocytosis is an innate mechanism of cells that is active all time, it is possible that the endocytosis events related to mRNA delivery are overlain by the endocytotic activity which always take place. Beside pH sensitive markers like the tested ones, the use of self-quenching markers could be suitable to record the events of interest. Markers like calcein are loaded in endosomes in a sufficiently high concentration that self-quenching of the dyes occur [161], which would result in a low fluorescence level if endosomes are intact. Calcein for example is an impermeable fluorescent dye, therefore it can cross the cell membrane only via endocytosis and is not able to lead to endsomal escape by itself [213]. In a transfection study using pDNA-polyplexes, calcein was used to determine the escape rate by detecting the two phenotypes of spatial signal distribution (punctuate pattern of intact endosomes and a uniform pattern after endosome escape), apparently the concentration used in the study was not high enough for self-quenching [213]. In this case the readout of the mean intensity time course would not show the excape event as the intensity of the whole cell is constant for both spatial phenotypes. If the concnetration of calcein is in the range of self-quenching and leakage of the endosomal membrane occurs, the intensity within the cell is supposed to increase because the marker is diluted. The dilution of the marker molecules means that they cannot quench each other anymore.

However, the use of self-quenching markers has to be tested carefully that high marker concentrations do not lead to side effects. It needs to be verified that a change of the marker signal occurs at the same time as mRNA release, as leakage of the membrane not necessarily relate to the release of macromolecules like mRNA [161].

7.3. Labeling transfected mRNA

The direct measurement of cytosolic mRNA concentration after transfection with the readout of fluorescent reporter protein expression time courses has the potential to improve the single-cell translation assay for several reasons. The parameter estimation would be based on two observables, the mRNA and protein concentration per time point, which is beneficial for the model fitting due to a reduced number of free parameters. Furthermore, by measuring the fluorescence intensity of cytosolic, labeled mRNA, which is proportional to the mRNA abundance, the number of released molecules m_0 can be determined. Thereby, the parameter non-identifiability of released mRNA molecules with the translation rate is resolved and the translation rate k_{TL} can be estimated independently. Figure 7.4. illustrates the expected cytosoloic mRNA dynamic of fluorescently labeled mRNA after transfection. The mRNA release event after endocytotic uptake should result in an sharp increase of the mRNA fluorescence intensity. The intensity maximum is proportional to the number of released mRNAs. After reaching the maximal value, the decrease in intensity should be describable by an exponential decay under the assumption that the mRNA decay is still conform to a first order kinetics. Of course, models assuming an enzymatic degradation could still be used and would benefit from two observables instead of one (see models in the Appendix B). The reporter protein expression e.g. GFP is expected to be delayed due to translation initiation and protein maturation. The GFP kinetics remain dependent on parameters like the translation rate or degradation rates.

In order to quantify the mRNA level by measuring a fluorescence intensity the mRNA needs to be labeled. For this purpose, several possibilities are imaginable and visualized in Figure 7.5.. The simplest approach is to directly label the IVT mRNA using nucleotides with a fluorescent dye covalently bound during synthesis. This approach results in synthetic mRNAs with several dyes per molecule (Figure 7.5. A), which is an advantage for detecting the fluorescence as several dyes emit more photons per molecule. The results of labeled mRNA time courses presented in subsection 7.3.1 are obtained using commercial Cy5 labeled mRNA encoding for eGFP (Trilink Biotechnologies). As an alternative to randomly distributed dyes within the whole mRNA sequence, a new method to label only the poly(A) tail was published recently [214]. Anhäuser et al. used IVT mRNA without an encoded poly(A) tail. They synthesized the poly(A) tail using yeast poly(A) polymerase and incorparted azido-modified adenosines. A fluorescent dye is bound to the modified adenosines using a bioorthogonal click reaction. The poly(A) labeled mRNA has the advantage that it can be excluded that the conjugated dyes lead to changes of the secondary structure within the UTRs. This can be important for the stability, or cause steric hindrances deteriorating protein binding of e.g. ribosomes, which could reduce translation efficiency dramatically. However, Althäuser et al. showed that the translation efficiency

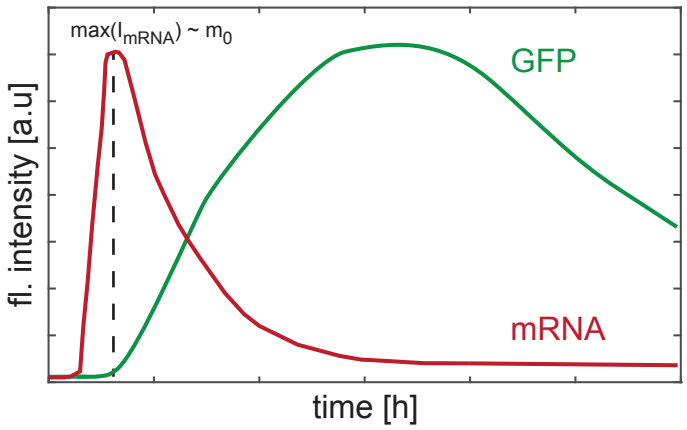

Figure 7.4.: Expected cytosolic mRNA concentration changes after transfection. The mRNA (red) shows a sharp increase shortly after transfection and a slow decrease over time. The maximal fluorescence intensity is proportional to the number of released molecules m_0. The protein expression level increases postponed to mRNA release and the dynamic depends on parameters like the translation rate or degradation rates.

of poly(A) labeled mRNA still leads to changes of the translation efficiency albeit to an improvement.

The other three possibilities which are discussed in the following are based on non-covalent labeling of mRNA. One common technique to label mRNA in living cells, which is often used in transcription studies is the MS2 system [215, 216, 217]. The MS2 coat protein, naturally occurring in bacteriophages, binds specifically to a RNA stem loop sequence [218]. The MS2 RNA-binding sequence needs to be cloned in the mRNA sequence and is typically repeated several times to enable multiple binding sites as visualized in Figure 7.5. B. As the MS2 protein is non-fluorescent by itself, a fluorescence (FL) protein is fused to the MS2 protein. In order to make use of the MS2 system in living cells, the cells of interest have to stably express the FL-MS2 protein and the sequence of the mRNA of interest need to contain the repeated stem loop sequence. For transcription studies, these results in the necessity of genetically engineered cells for the expression of the artificial FL-MS2 protein as well as the altered mRNA sequence. For translation studies after mRNA transfection the genome of the host cell still needs to be changed for stable expression of FL-MS2 proteins. The stem loop repeats need to be cloned in the pDNA vector used for IVT of the mRNA. Even ignoring the necessary genetic modification, the expression of FL-MS2 proteins leads to a permanent fluorescent background and single mRNA molecules can only be detected due to FL protein accumulation if several FL-MS2 proteins bind to one mRNA molecule.

In order to avoid the preparation of genetically engineered cells and the high fluorescent background, the use of molecular beacons is promising for detection of mRNAs after transfection. A molecular beacon (Figure 7.5. C) consists of a oligonucleotide, which forms a hairpin structure, with a fluorescent dye at one end and a quencher molecule at the other end [219]. This oligonucleotide design leads to only a weak fluorescence for the hairpin conformation as the quencher molecule prevents the emission of photons of the fluorescent

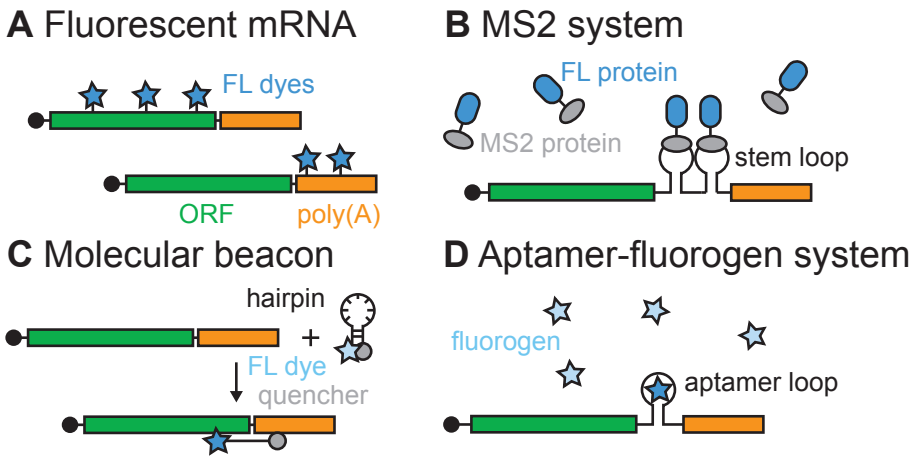

Figure 7.5.: mRNA labeling technologies for transfection studies. (A) Fluorescent mRNA: The mRNA is directly labeled during IVT by using nucleotides conjugated with fluorescent dyes. As alternative the poly(A) tail can be labeled exclusively, if the poly(A) tail is synthesized enzymatically after IVT using labeled adenosines. (B) MS2 system: labeling technique based on the interaction of the bacteriophage coating protein MS2 with a stem loop RNA sequence. A fluorescent protein is fused to the MS2 protein. (C) Molecular beacon: A oligonucleotide forming a hairpin is conjugated with a fluorescent dye at one end and a quencher molecule on the other side resulting in a molecular beacon. The oligonucleotide sequence is able to hybridize to the mRNA and the dye is not quenched any more and therefore fluorescent. (D) Aptamer-fluorogen system: A RNA aptamer loop binds a small peptide (fluorogen) and induces the fluorescence of that fluorogen.

dye. In order to label a mRNA of interest a part of the oligonucleotide sequence needs to be complementary to the mRNA sequence. If mRNA and the molecular beacon are in close proximity, the oligonucleotide hybridizes to the mRNA by Watson-Crick base pairing as the duplex formation of mRNA and molecular beacon is more stable than the self-interaction of the molecular beacon. If the quencher molecule and the fluorescent dye lose their spacial proximity, the fluorescent dye is able to emit photons again if excited by a suitable excitation wavelength. In order to enhance the number of emitted photons per molecule, several molecular beacons can be engineered binding to different regions of the mRNA. The specificity of the fluorescence signal can be further improved by using a dual molecular beacon approach based on Förster resonance energy transfer (FRET) [220]. Two groups of molecular beacons are designed having different fluorescent dyes at one end and a suitable quencher molecule at the other end of the oligonucleotide. The two dyes need to be a suitable FRET pair of a donor and an acceptor chromophore. If the donor and the acceptor molecule are in close spatial proximity without being quenched, the excited donor chromophore can transfer energy radiation free on the acceptor molecule. This results in an excited acceptor molecule that returns into the ground state through the emission of a photon. The use of molecular beacons has the advantage that the cells, which will be transfected, as well as the IVT transcribed mRNA can be used without any

changes. The molecular beacons need to be designed and synthesized depending on the mRNA of interest taking inter- and intramolecular interaction as well as specificity into account. The molecular beacons need to be transfected besides the exogenous mRNA. For single-cell translation studies it is necessary to adapt the transfection protocol to ensure that enough molecular beacons are present to label all of the released mRNAs and to avoid toxic effects due to double transfection. Furthermore, it is possible that the binding of the molecular beacons leads to steric hindrance of translation related proteins dimishing the translation efficiency.

Besides the use of molecular beacons, the approach of RNA mimics of GFP is promising for single-cell translation studies. The approach is based on a synthetic analog of the GFP fluorophor to fluorescnetly label mRNA since GFP is used to label proteins of interest [221]. The HBI fluorophor structure of GFP (as described in section 2.1.) was the basis to develop a fluorogen, which processes a high quantum yield if bound to a RNA aptamer as illustrated in Figure 7.5. D. Several derivates of HBI were analyzed and none of them showed detectable fluorescence intensities within cells. This means that the fluorogen is most likely not activated by any naturally occurring RNA or DNA structure. Suitable aptamer sequences were screened based on a RNA molecule library of $5{\cdot}10^{13}$ candidates and the interaction with the HBI derivates was analyzed and optimized regarding the quantum yield of the aptamer-fluorogen combination [52]. Paige et al. developed a RNA-fluorogen couple consisting of 3,5-difluoro-HBI (DFHBI) binding to one specific RNA sequence called "Spinach" [52]. Spinach has a quantum yield of 0.72, which is 20% higher than the one of eGFP. However, the molar brightness of Spinach is only about 53% of the value of eGFP [52]. As the concentration of transfected mRNAs is much lower than the concentration of expressed eGFP, the fluorescence intensities measured per cell are expected to be much lower for Spinach compared to eGFP. But, an advantage of Spinach is low photobleaching compared to eGFP. This effect is presumably explained by the exchange of unbound and bound DFHBI minimizing the accumulation of photobleached complexes [52]. In order to obtain a higher fluorescence signal per molecule Zhang et al. showed that the fluorescence intensity increases linearly with the number of Spinach repeats per mRNA without influencing the translation efficiency or mRNA degradation in bacteria [222]. In order to measure mRNA and FL kinetics within one cell a variety of aptamer-fluorogen systems are available covering a broad visible spectrum [221, 223]. The same is the case for FL proteins resulting in a large number of possible combinations of aptamer-fluorogen system with a FL protein, which are spectrally distinguishable from each other.

Each of the mentioned approaches has their advantages and drawbacks. Several requirements must be considered, in order to decide which approach is the most promising to quantify the mRNA dynamic after transfection at the single-cell level. Firstly, the signal has to be dependent on the functionality of the mRNA. If a mRNA is degraded the fluorescence signal needs to break down. This is not the case for fluorescent nucleotides incorporated in mRNA molecules, as will be described in subsection 7.3.1. in detail. It is expected that the other three approaches lead to decrease of the fluorescence signal if the mRNA is degraded. The mRNA sequence domains are destroyed by degradation so that fluorescence cannot occur anymore. Secondly, as a low number of mRNA molecules is

expected inside the cell the signal needs to be as bright as possible. As the FL-MS2 proteins are fluorescent all the time the molecule is only detected by accumulation of several dyes due to the repeated hair pin structures. If the mRNA intensity is readout as mean intensity of the cell analogous to eGFP no dynamics will be visible as the total amount of FL-MS2 proteins is independent of the mRNA transfection. Furthermore, high resolution microscopy is needed to detect the spots of accumulated FL-MS2 proteins, which would minimize the throughput of the assay. For the molecular beacon as well as the aptamer-fluorogen approach a very low fluorescence background is expected, as the fluorescence of the molecular beacons is quenched in the hairpin conformation and the fluorogen itself is non fluorescent if not bound to the appropriate RNA aptamer. Thirdly, the mRNA properties such as the translation rate should be influenced as little as possible. As it is likely that the hybridization of several oligonucleotides to a mRNA molecule affects the binding of proteins related to translation or degradation the properties and therefore the translation dynamics are altered for the molecular beacon method. Using a tandem Spinach array these processes seem to be not affected in bacteria [222]. Considering these aspects one comes to the following conclusion. A tandem Spinach array cloned in the IVT mRNA could be suitable to detect a sufficient mRNA fluorescence in single-cells to quantify mRNA dynamics in mammalian cells. Yet, it needs to be clarified if the mRNA signal is cytosolic or if mRNAs trapped in endosomes are labeled as well. This would lead to a dynamic which is a superposition of the cytosolic and endosomal mRNA dynamic. The mRNAs are in different environments, such as pH, so different degradation kinetics might occur. However, the fluorescence intensity of Spinach is known to be pH dependent and is reduced at low pH compared to physiological pH [52]. For the possible superposition of several mRNA kinetics a live-cell imaging method for transfected mRNA needs to be established taking many aspects into account.

7.3.1. Fluorescent dyes conjugated to nucleotides

The use of directly labeled mRNA may not give information on the mRNA degradation kinetics but it may still enable the labeling of transfection complexes and the measurement of uptake kinetics in single-cells. The results presented in this subsection were obtained using the single-cell translation assay as described previously using eGFP coding mRNA of which 10% of mRNAs is directly labeled with Cy5 flurophores (Trilink Biotechnologies). During the time-lapse measurement the signals of Cy5 and eGFP were recorded simultaneously resulting in two image stacks. These image stacks were overlaid and both mean intensity time courses for each transfected cell were read out using the MWA plug in. The reason why only a fraction of 10% labeled mRNA was used is that cells transfected with 100% labeled mRNA showed a dramatically decreased eGFP expression compared to cells transfected with the same amount of unlabeled mRNA. It was concluded that the fluorescent dyes lead to a decrease of the translation efficiency and only the usage of a small fraction of labeled mRNA enabled to still measure the eGFP expression kinetic.

Lipoplex

The use of labeled mRNA encapsulated in L2000 lipoplexes enabled a visualization of the lipoplexes like shown in Figure 7.6. A. The L2000 lipoplexes (red) are randomly distributed during complex incubation (Figure 7.6. A left image). The yellow outlines mark the positions of the FN squares. The location of the lipoplexes did not allow conclusions to be drawn if a FN square was occupied by a cell or not. After one hour of incubation, the unbound lipoplexes were washed out using the perfusion system. After the washing step the number of lipoplexes is obviously reduced and most of the remaining lipoplexes are located on FN squares, which were occupied by cells (Figure 7.6. A right image).

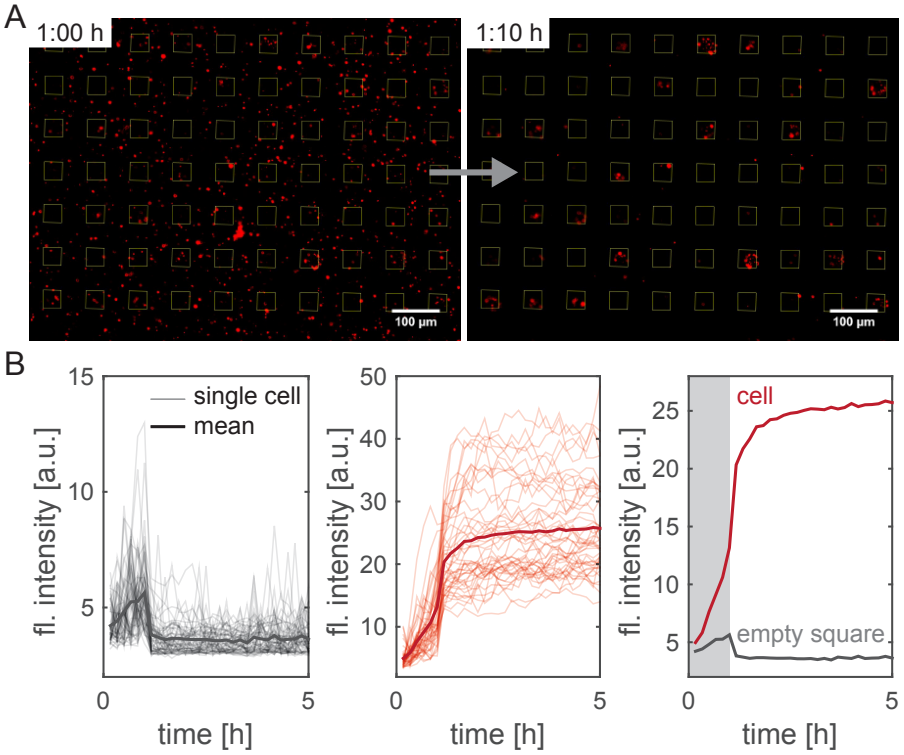

Figure 7.6.: Cy5 labeled mRNA encapsulated in L2000 lipoplexes. (A) The L2000 lipoplexes are randomly distributed over the field of view (left image). After 1 h of incubation the unbound complexes are flushed out and the remaining complexes are mainly adsorbed on cell surfaces (right image). The yellow outlines mark the FN square pattern. (B) The fluorescence intensity signal of the Cy5 labeled mRNA of empty squares shows a small signal increase during complex incubation and an abrupt signal drop after the washing step at t=1 h (first plot). The signal of cell occupied squares shows a faster and more pronounced increase compared to empty squares saturating after about 2 h (second plot). The mean signals of empty (gray) and occupied squares (red) illustrate the complex adsorption on the cell surface (third plot). The gray bar indicates the transfection incubation period. Adapted from the supplementary information of [P3].

The Cy5 signal of empty squares showed an increase during the first hour of lipoplex incu-

bation due to adsorbing lipoplexes on the surface. An abrupt decrease to the background level after flushing the only partially bound lipoplexes (Figure 7.6. B first plot). The Cy5 signal of lipoplexes adsorbed on cells on the other hand showed an increasing signal even after the washing step that saturates after about two hours (Figure 7.6. B middle plot). The mean trajectories of empty squares (gray) and cells (red) illustrates the different temporal behavior (Figure 7.6. B last plot). During lipoplex incubation the signal of Cy5 increased more than that of empty squares and interestingly increased even further for another hour. This can be interpreted that lipoplexes adsorb strongly on the cell surface and rarely detach, which can happen on empty squares. Furthermore, it is likely that the Cy5 fluorescence of labeled mRNAs increases during lipoplex uptake because of structural remodeling of the lipoplexes inside the endosomes leading to dequenching of the Cy5 fluorescence, which occurs if the mRNA is densely packed inside the lipoplexes [140].

Lipid nanoparticle

Analog experiments to L2000 lipoplexes with 10% Cy5 labeled mRNA encapsulated were performed using the three LNP formulations DLin, KC2, and MC3 (section 5.5.). Representative images of one time-lapse stack recording HuH7 cells transfected with MC3 LNPs are shown in Figure 7.7.. The fluorescence signal of the Cy5 labeled mRNA (red) was overlaid with eGFP (green) for different time points of the measurement. The time indicated in the upper left corners refers to the addition of the MC3 LNPs to the cells. After 30 min the cells aligned on the micropattern became visible as a mRNA signal appeared. Interestingly, randomly distributed complexes as observed for the L2000-lipoplexes (Figure 7.6. A) could not be monitored. After stopping the LNP incubation by flushing unbound LNPs after one hour the first cells already showed eGFP expression, which is in agreement with the onset time distribution (Figure 5.10. A). After 2 h most of the cells of the representative field of view showed eGFP expression and the expression level clearly increased over the next hours as can be seen for the overlay at 10 h after transfection.

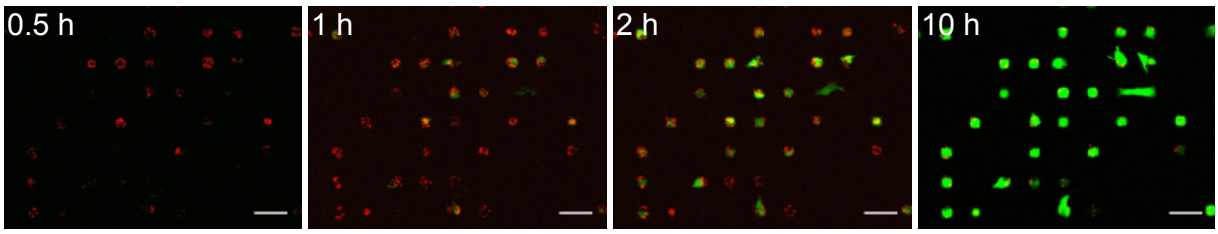

Figure 7.7.: Overlaid fluorescence images of Cy5 labeled mRNA (red) and eGFP (green) after transfection using MC3 LNPs. The time is relative to the addition of the LNP particles and the scale bars correspond to 100 µm.

The fluorescence time courses of Cy5 and eGFP were readout using the MWA plugin. Representative time courses are shown in Figure 7.8. A. The single-cell time courses of the Cy5 signal showed almost no increase during the first 20-30 min after adding the LNPs (indicated by the first arrow). This is different compared to the Cy5 time courses

of L2000 transfected cells (Figure 7.6. B). For some cells, like the red time course, the signal increased after the incubation time (gray bar), which was not observed for L2000 transfected cells. But, comparable to cells transfected with L2000 the Cy5 signals saturated after roughly 2 h. The different kinetics for Cy5 labeled mRNA transfected with LNPs can probably be attributed to different quenching behavior of Cy5 mRNA encapsulated in LNPs compared to L2000 lipoplexes. The mRNA encapsulated in LNPs is expected to be denser packed in LNPs than in lipoplexes. It was shown for siRNA containing LNPs that LNPs show a high electron density of the core, which is in contrast to the less dense structure of lipid vesicles [182]. Probably, the Cy5 fluorescence is almost completely quenched for intact mRNA-LNP systems and the measured time courses correspond to the kinetic of structural remodeling of the LNPs during particle uptake.

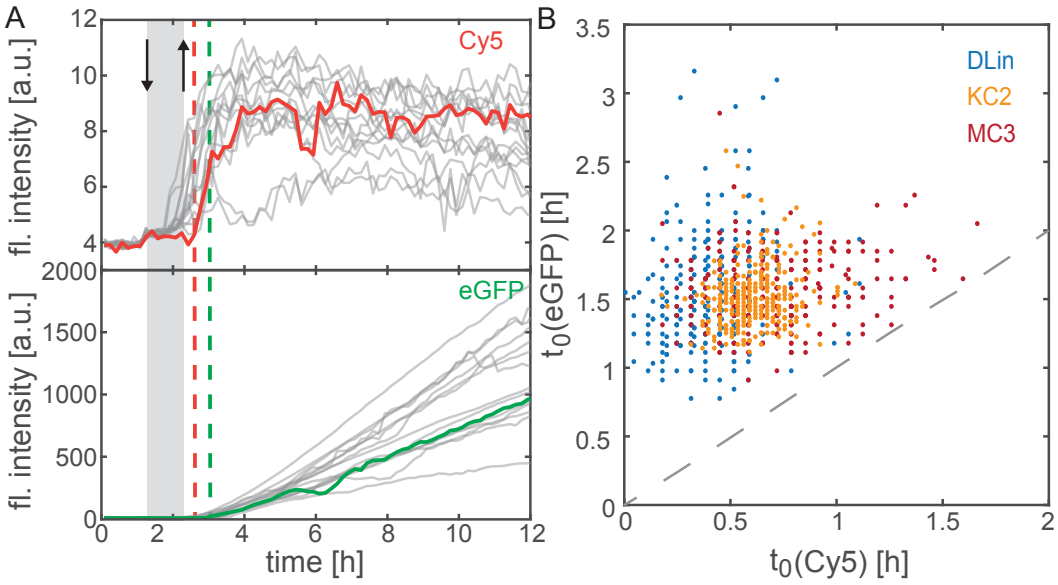

Figure 7.8.: Two signal readouts of Cy5 labeled mRNA and expressed eGFP at the single-cell level. (A) The fluorescence trajectories of Cy5 labeled mRNA is shown together with the corresponding eGFP signals. The gray marked area represents the 1 h incubation time of the LNPs. The arrows indicate the injection of the LNPs diluted in cell growth medium containing 1% FCS and the washing step with cell growth medium containing 10% FCS after one hour. The representative colored time courses were measured within the same cell. The dashed lines mark the onset time point of the Cy5 signal (red) and the eGFP signal (green). (B) Scatter plot of the Cy5 onset t_0(Cy5) vs. the eGFP onset t_0(eGFP) for cells transfected with DLin (blue), KC2 (orange), or MC3 (red). Each dot corresponds to a single cell.

Due to the simultaneously measurement of mRNA and eGFP signal, correlation studies of both onset time points were possible. The onset time points of Cy5 t_0(Cy5) and eGFP t_0(eGFP) were determined manually for the results presented in Figure 7.8. B. The Cy5 time courses were very noise and therefore more difficult to determine automatically. Subsets of the data discussed in section 5.5. of cells transfected with one of the three investigated LNP formulations (DLin, KC2, and MC3) were used for the correlation study.

The mean onset times for eGFP expression of all LNP formulations are in the same range than the ones determined by fitting the translation-maturation model (manual: ∼1.5 h, fitting: ∼1.4 h). The mean onset time of Cy5 differs between the formulations (DLin: 0.40 ±0.20 h, KC2: 0.60 ±0.14 h, MC3: 0.72 ±0.28 h). The onset times for Cy5 and eGFP are uncorrelated for all three LNPs with Pearson's correlation coefficients of 0.11 for DLin, 0.18 for KC2, and 0.13 or MC3 (with p-values lower than 0.05 for all data sets). Due to the variation of the mean Cy5 onset times, the mean duration between the Cy5 onset and the eGFP expression is different as well. The mean time-lags are 1.14 ±0.41 h for DLin, 0.91 ±0.23 h for KC2, and 0.85 ±0.36 h for MC3. Interestingly, this is of the same order than the pK_a values (DLin: 6.8, KC2: 6.7, MC3: 6.4). As the onset time points are determined manually, the results need further verification to exclude a bias e.g. by adapting the clustering approach introduced in subsection 5.3.2. for the Cy5 time courses.

As expected the directly labeled mRNAs are not a suitable approach to detect mRNA degradation in living cells or to determine the number of released mRNAs. Since the Cy5 signal did not decrease over a time period of 10 h, it can be assumed that the Cy5 dyes are not affected by degradation. During a 10 h time period and a mean mRNA half-life of 13.4 h (as determined in subsection 5.4.1.) a decrease of 40% is expected for a exponential degradation of the successfully transfected mRNAs, which was not observed for the Cy5 time courses shown in Figure 7.8. A. Nevertheless, the Cy5 time courses of cells transfected with Cy5 labeled mRNA using L2000 lipoplexes and LNPs showed partly different kinetics. This can lead to the conclusion that the labeled mRNA encapsulated in LNPs may give temporal information on the endocytosis process and the structural changes of LNPs during endocytosis.

7.3.2. Fluorescence *in situ* hybridization to label exogenous mRNA

The establishment of a live-cell mRNA labeling method does not appear to be trivial. However, for fixed cells there is a widely used approach called fluorescence *in situ* hybridization (FISH). The method is mainly used for transcription studies to quantify the number of mRNA molecules of a certain gene at the time point of cell fixation in order to calculate transcription rates [62, 224, 225, 226]. The idea of labeling is similar to the molecular beacon approach (Figure 7.5. C). A set of fluorescently labeled oligonucleotides of ∼20 nt lengths with complementary sequences to the targeted mRNA needs to be designed accordingly and synthesized. The cells are fixed using e.g. formaldehyde and the plasmamebrane is permeabilized. The oligonucleotides are added in an appropriate buffer to the cells and incubated for at least a few hours. During incubation the oligonucleotides diffuse into the cells and are able to hybridize to the mRNA of interest. The technique is capable of achieving single mRNA resolution as up to 50 fluorescent dyes, mark one molecule depending on the number of different oligonucleotide sequences [227]. After washing the unbound oligonucleotides to reduce the background signal to a minimum, the cells are ready to be imaged using higher magnification (60x and higher). In order to record the fluorescence intensities of the labeled mRNAs, a z-stack per field of view is recorded to cover the whole cell volume. These z-stacks are further used for image processing to read out the number of mRNA spots and their intensities.

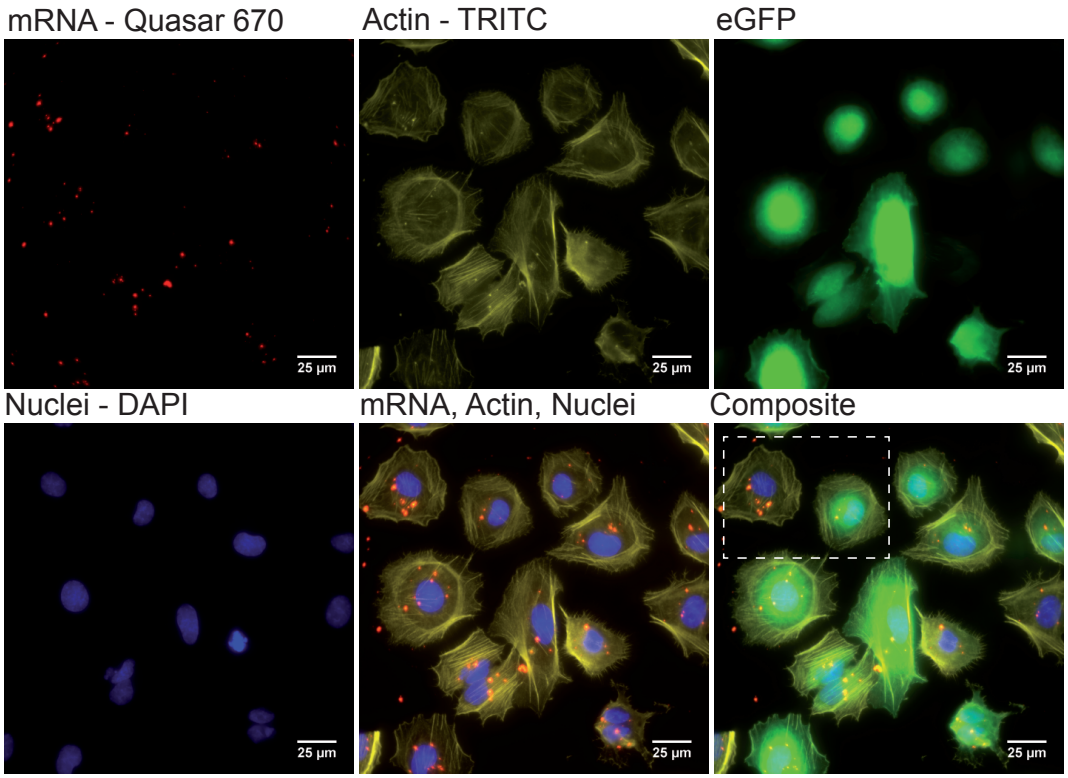

Figure 7.9.: Maximal z-stack intensity projections of HuH7 cells encoding for eGFP 20 h after mRNA transfection. The mRNA is labeled using FISH probes labeled with Quasar 670 (red). For cell recognition the actin skeleton was stained with Phalloidin-TRITC (yellow). The nuclei were stained using DAPI (blue). The expressed eGFP after transfection is colored in green. The overlay of the mRNA, actin, and nuclei signal illustrates the spatial mRNA distribution. The overlay of all four recorded signals shows that the presence of mRNA does not necessarily lead to eGFP expression as highlighted by the dashed rectangle. The scale bars correspond to 25 µm.

In order to establish a FISH protocol, HuH7 cells were seeded in six-channel slides and transfected with mRNA encoding for eGFP 4 h after seeding. The cells were fixed with formaldehyde 20 h after transfection. Afterwards, the labeled oligonucleotides recognizing the eGFP coding sequence were incubated with the cells over night to enable hybridization to the transfected mRNA. A detailed protocol for FISH is given in Appendix A. Figure 7.9. shows a representative z-stack projection of cells with fluorescently labeled mRNAs using the established FISH protocol 20 h after transfection. The labeled mRNA spots (red) appeared to be randomly distributed over the field of view. For easier localization of the spots, if they are intra- or extracellular, the cells were stained with DAPI to label the nuclei (blue) and Phalloidin-TRITC to make the actin skeleton (yellow) visible. Many of the cells show a strong cytosolic eGFP signal as the cells had several hours to express the protein before fixation. The overlay for the mRNA, actin, and nuclei signals illustrates that most of the mRNA spots are located inside the cells as most of the mRNA complexes are flushed after transfection incubation. The other overlay of all four signals

exemplifies that a cell with labeled mRNA did not necessarily express eGFP, whereas the neighboring cell showed fewer mRNA spots but eGFP expression (dashed rectangle). The mRNA spots of the non fluorescent cell for eGFP are most likely trapped in endosomes or still bound on the cell surface without sufficient release of mRNA molecules in the cytosol. Interestingly, the mRNA entrapped in endosomes and the nanocarrier are not degraded (completely) 20 h after treatment otherwise the oligonucleotides could not have hybridized.

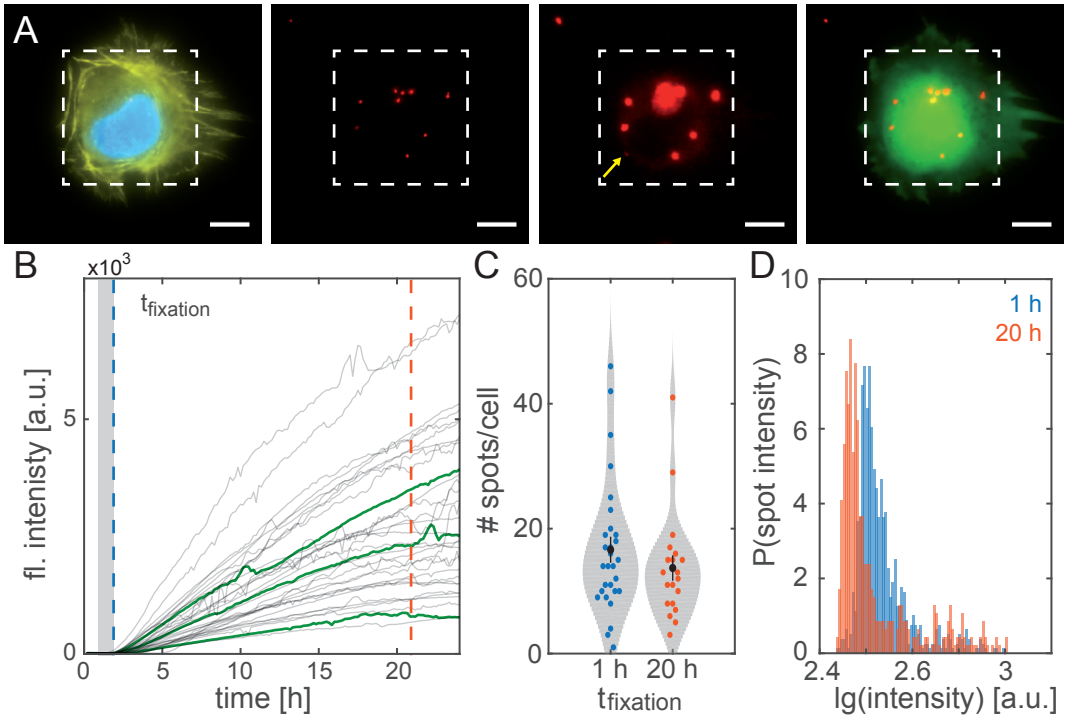

Figure 7.10.: mRNA labeling using FISH combined with single-cell translation assay. (A) Representative cell on FN square (dashed outline). The nucleus (blue) and the actin skeleton (yellow) of the cell are stained. The red spots correspond to mRNA clusters labeled with Quasar670 dyes at two different contrast settings. The green signal correspond to eGFP fluorescence. All images are maximal intensity projections of a z-stack and the scale bars correspond to 10 µm. (B) Single-cell expression time courses obtained parallel to FISH measurement. The gray bar indicates the transfection period and the dashed lines mark the times of cell fixation for FISH labeling. (C) Number of mRNA spots determined per cell at two different fixation times. (D) Histograms of the spot intensities 1 h (blue) and 20 h (orange) after transfection.

The established FISH protocol was further combined with the single-cell translation assay to determine the number of mRNA molecules at different time points of the time-lapse measurement. All six channels of the single-cell array were transfected under the same conditions and the eGFP fluorescence was recorded during the time-lapse measurement. The channels of the array were fixed by flushing formaldehyde at different time points in order to fix the cells in one channel. Figure 7.10. A shows a representative cell on a FN square (dashed outline). The nucleus (blue) and the actin skeleton (yellow) illustrate the

cell position and shape. The fluorescence image of the labeled mRNA is shown for two different contrast settings. At high contrast, sharp spots are visible. If the contrast is adapted, spots of weaker intensities are visible (marked by the yellow arrow), which are almost outshined by the bright spots. The eGFP signal of the cell is distributed over the whole cytoplasm, as expected. In Figure 7.10. B single-cell time courses of eGFP expression are shown for cells fixed at the end of the time-lapse measurement. The gray bar indicates the incubation period of mRNA transfection. The dashed lines mark the time points of fixation at 1 h (blue), directly after transfection incubation, and at 20 h (orange) after transfection. The number of mRNA spots was determined for a small number of cells (Figure 7.10. C) for both fixation times $t_{fixation}$. The number of spots per cell appeared to decrease between 1 h and 20 h, but did not show a large variation. As the two contrast settings for labeled mRNA already indicated, it is unlikely that each spot correspond to one molecule due to the high variation in intensity. This assumption is often made for mRNA detection studies after transcription as it is unlikely that the mRNAs form clusters [224]. For mRNA after transfection it is expected that clusters are visible because the mRNA was complexed in nanoparticles or trapped in endosomes. Furthermore, it is not known if the released mRNAs after endosomal escape are freely diffusing as single molecules in the cytosol. In order to approximate the total number of mRNAs shifts, the intensity distribution of the mRNA spots was determined. It is expected that the distribution shifts to smaller intensities over time due to mRNA degradation. In Figure 7.10. D the distribution of spot intensities shows a shift between $t_{fixation}$ of 1 h and 20 h as expected.

In conclusion, the experimental FISH protocol to fluorescently label mRNA was straight forward to establish. However, the image processing of exogenous mRNA clusters appeared to be more complicated than for endogenous mRNA in transcription studies. For FISH image stacks of endogenous mRNA, image processing software is available that allows for quantification [228]. This software was not applicable for the presented data due to the larger range of spot intensities of exogenous mRNA clusters. Kirschman et al. recently used FISH to study mRNA delivery in order to correlate the mRNA to the protein level at the single-cell level after transfection. They reported the same difficulties due to mRNA cluster formation [136]. They excluded mRNA clusters which colocalized with a endosome marker and determined these cluster as non functional and only took the cytosolic mRNA clusters into account for correlation studies. If FISH should be used to determine the number of transfected molecules and for mRNA-protein correlation studies, further refinements of the protocol are needed especially for the image analysis.

7.4. Discussion

In this chapter, three approaches to take advantage of two fluorescent markers for single-cell translation studies were presented. The measurement of two signals within the same cells enables a joint analysis of both fluorescence intensity time courses. Firstly, the expression of two reporter proteins was discussed. By co-transfection of the genetic code for both reporters, a large fraction of cells can be observed that express both proteins. The fluorescent protein expression time courses can be read out analogously to a single

one. This approach was applied for siRNA silencing studies at the single-cell level [M4] as one mRNA was targeted by siRNA and the expression of the second mRNA was used as an intracellular reference (Figure 7.1.). The approach could be even further improved by the synthesis of mRNA constructs encoding for both proteins at the same time. The expression kinetics for both proteins would depend on the same number of transfected mRNAs and the same mRNA degradation rate. This could be achieved by mRNA sequences containing both protein coding sequences that have a so called "2A" box in between the coding sequences. The 2A box encodes for a short peptide which is enzymatically cleaved after translation resulting in two separated proteins [229, 230]. For quantitative analysis the expression of both proteins could be modeled with one set of ODEs taking the cleavage efficiency into account and that two proteins are translated by the same mRNA. This would reduce the number of parameters compared to model each reporter expression kinetic by an identical ODE model with varied parameters for e.g the degradation rates.

Secondly, the fluorescence intensity time courses of two endosome markers were analyzed at the single-cell level to screen if they show characteristic events that could be correlated to mRNA delivery. As endocytosis is an innate mechanism of the cell and cells engulf particles and fluids from the environment all the time, the uptake of mRNA nanoparticles is likely to be superimposed by the remaining endosomal activity. However, it is possible that the endosomal escape events can be detected as the endosomes do not leak or rupture normally [213]. Yet, such events were not detected by the two tested markers, pHodro and Lysotracker, as can be seen in the respective fluorescence time courses (pHodro: Figure 7.2., Lysotracker: Figure 7.3.).

Finally, several options for labeling mRNA were discussed. The most promising methods for live-cell labeling of exogenous mRNA are the molecular beacon and Spinach approach. But, both need careful implementation that could be done in the future. A simple method to visualize mRNA, is the usage of mRNA synthesized with fluorescently labeled nucleotides e.g. Cy5. The labeled mRNA was encapsulated in L2000-lipoplexes (Figure 7.6.) or LNPs (Figure 7.7.) and the fluorescence signal for mRNA-Cy5 and eGFP was recorded in parallel. The time courses of labeled mRNA encapsulated in LNPs enabled the correlation of the mRNA onset time with the eGFP onset time (Figure 7.8.). It was shown that, as expected, the eGFP signals is always delayed to the one of Cy5 by roughly one hour. Interestingly, the two event times are still uncorrelated (Pearson's coefficient lower than 0.2 for all three investigated LNP formulations). The use of directly labeled mRNA has two major drawbacks. Firstly, the labeled mRNA is not translated as well as unlabeled one and therefore only a fraction of 10% labeled mRNA was transfected together with unlabeled mRNA. Secondly, the signal is independent on mRNA degradation as the dyes remain fluorescent after degradation of the mRNA. In order to determine the number of transfected mRNAs, the mRNAs after transfection were labeled using FISH. The method requires cell fixation and therefore loses time resolution. But, it was directly shown that the number of mRNA clusters has a broad distribution, which shows only small differences between an early fixation after transfection and a late fixation. However, the intensity distributions of the spots, which is expected to be proportional to the number of mRNA molecules, change depending on time. The results were explained

due to the fact that the majority of transfected mRNAs are trapped in endosomes, as it was shown for siRNA [135] and mRNA [136].

In summary, the gain of information through the single-cell translation assay benefits from the application of a dual-color approach. Yet, the choice of the second marker needs to be checked accurately to ensure that the marker does not have any side effects and is specific for the feature of interest e.g. endosome rupture or cytosolic mRNA. Ideally, a fluorescent labeling technique needs to be developed which is specific for intact, cytosolic mRNA and suitable for live-cell imaging. This would allow to record the mRNA and protein kinetics in parallel for a combinatory dynamic analysis. This method could possibly further be combined with a marker for endosomal escape, which would lead to an even more accurate resolution of the mRNA transport and expression process.

8. Conclusion and future prospects

In this thesis, the quantification of mRNA nanoparticle transport and transfection mediated protein translation was obtained by analyzing single-cell fluorescence intensity time courses. The single-cell parameters were obtained from expression kinetics of fluorescent reporters. To this end, deterministic rate equations were formulated in order to describe the complex biological network within a mathematical framework. The population distributions of the expression rates as well as the degradation rates for mRNA and protein were estimated. The parameter set for each single cell enabled quantitative studies of the parameters depending on different mRNA transport conditions or varied mRNA sequences. Figure 8.1. illustrates that single-cell microarrays enable the quantification of mRNA transport and translation kinetics.

Thereby, it was revealed that the delivery time of mRNA nanoparticles does not correlate with the expression efficiency at the single-cell level. However, the distributions showed systematic changes dependent on extracellular factors like the nanoparticle formulation and serum protein adsorption (chapter 5). Single-cell expression onset time distributions after mRNA transfection, were measured here for the first time, which revealed that the delivery process occurs over a time period of only a few hours. The method enables new insights into the correlation of nanocarrier delivery timing and efficiency. The fact that these parameters are not correlated within single cells indicates that there is no intrinsic mechanism that links the transport timing to the total protein expression rate. Additional efforts are needed to resolve the intermediate transport processes of nanoparticle delivery to gain a better understanding of which intermediate steps are crucial for effective delivery. This could possibly be achieved by endocytotic markers like discussed in the previous chapter.

Beside the study of mRNA nanoparticle transport, the assay offered the chance to determine protein as well as mRNA half-life distributions depending on mRNA sequence variations on the single-cell level. An improved understanding of the expression kinetics and how the expression can be controlled is important for successful treatment design in gene therapeutic applications. Furthermore, the expression of fluorescence proteins with distinct half-life distribution caused by different coding sequences enabled the integration of perturbation experiments which led to a more robust parameter estimation (chapter 4). The newly developed modeling approach based on multi-experiment nonlinear mixed effect modeling (NLME) gave even insight in the nature of batch effects, which most likely occur due to a high variation of the average number of transfected mRNAs for different experimental replicates. The characterization of sources for experimental day-to-day variances could be used for further improvements of the experimental set up and therefore data quality.

As the regulation of mRNA degradation is a key mechanism for post-transcriptional gene

regulation, the dependency of the poly(A) tail length on the mRNA stability was quantitatively analyzed at the single-cell level. The population distribution of the mRNA half-life for constructs of ten different poly(A) tail lengths were measured (chapter 6). A model based on a power law function was introduced that described the experimental data for all investigated poly(A) tail lengths. It was shown that the mRNA half-life as a function of poly(A) tail length is described by a power law function with exponent 0.53. A critical length for the poly(A) tail was derived from the data. This length is needed to enable protein expression. Based on the proposed model, it was confirmed that the length necessary for binding of two PABPs is an important benchmark for mRNA stability. For future prospects, the measurement of a degradation dependent mRNA signal would allow to determine the translation rate as single parameter. Using the NLME approach by integrating the data sets of proteins expression with varied poly(A) tail length, it could be clarified whether translation rate and mRNA half-life correlate at the single-cell level and possibly more detailed insights in mRNA turnover could be achieved.

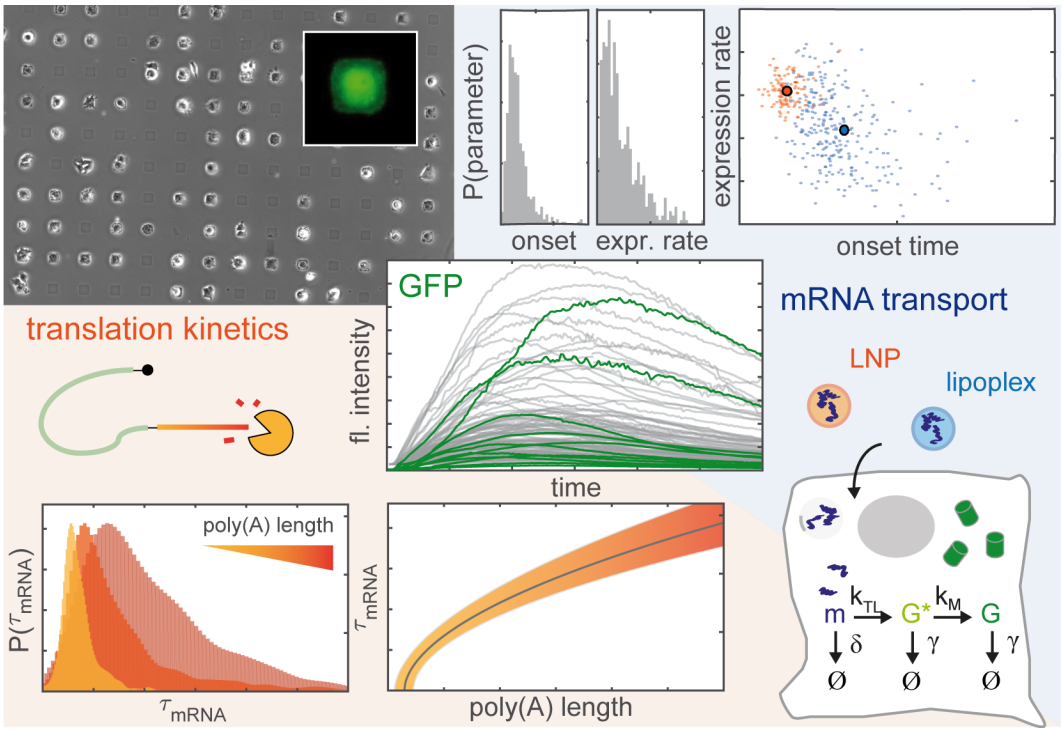

Figure 8.1.: Single-cell microarrays enable the high-throughput readout of fluorescence protein expression time courses mediated by mRNA transfection that can be modeled by biochemical rate equations. The method enables the systematic investigations of mRNA nanoparticle transport (blue) and translation kinetics (orange). The onset time-expression rate profiles of cells transfected with LNPs or lipolexes are shown as example for mRNA transport studies. As example for translation kinetic studies the dependency of mRNA stabiliy on poly(A) tail length is shown. The mRNA half-life distributions τ_{mRNA} increase for longer poly(A) tail lengths. The dependency of τ_{mRNA} on poly(A) tail length is described by a power law function.

Single-cell translation studies mediated by mRNA transfection show great potential as a novel method to quantify various aspects of the process. The studies showed, for instance, that the choice of micropattern was critical for the quality of the single-cell traces regarding to the measurement noise (chapter 3). The noise of the protein expression time courses were found to be autocorrelated. This finding could be used to study the cell confinement on microstructures in detail as it is likely that the autocorrelation occurs dependent on the formation of cellular protrusions. The µPIPP method was chosen for further studies as µPIPP arrays are easy to fabricate for small numbers of arrays. Compromises in reproducibility regarding day-to-day variances affecting the cell confinement had to be accepted. The development of new methods that enable the fabrication of single-cell arrays in a more robust fashion with better cell confinement would have positive effects on the noise of the single-cell time courses and thus on parameter estimation and on statistics as more cells could be analyzed. A promising method to produce single-cell arrays is by the use of photopatterning [231]. Currently, this method is established in our laboratory in cooperation with ibidi. The basic idea is to generate a micropattern of adhesion spots on a thin cell repellent hydrogel layer. The micropattern is obtained by a single illumination step through a mask followed by a click reaction to functionalize the microstructures with a RGD peptide to enable cell adhesion. The noise analysis of time courses dependent on different microarray fabrication methods revealed that the time courses not only depend on the microarray but also on the algorithms used for image processing. To date, the time needed to analyze the image stacks takes at least as much time as the experiment itself. New algorithms for image processing and a better automation of the fluorescence intensity read out would be beneficial for the assay. The image processing could, for instance, be automated by using published algorithms based on deep learning for cell or pattern recognition combined with a automated selection of single-cell time courses excluding unsuitable time courses of cells e.g. due to cell division [232]. The signal readout could be further extended by recognizing cell internal fluorescence differences like a punctuated fluorescence pattern e.g. for endosomes which probably change over time and gives insights in release kinetics.

Further, new insights on the cell-to-cell variability that arise in gene expression could be obtained by combining the quantification of translation kinetics for hundreds of cells with the global quantification of mRNA and protein abundance. The Single-cell translation assay enables high time resolution for one protein, whereas global quantification allows the measurement of thousands of mRNAs and proteins at a critical time point using e.g. single-cell RNA-Seq or mass spectrometry. Recently, Cheng et al. quantified the cell population response to misfolding stress on mRNA and protein expression in mammalian cells for four time points over a duration of 30 h [29]. They analyzed the regulation of mRNA and protein expression on a global level using transcriptomics and proteomics approaches. Kinetic parameters were estimated for synthesis and degradation for both mRNA and protein using a new statistical tool. It was shown that the two regulatory levels were equally important, but their impact on molecule abundance differed. Combining the approaches of such studies with the single-cell translation assay would allow investigating the protein expression of one reporter with much higher time resolution and even more important would give insights in the cell-to-cell variability. The misfolding stress was

induced by adding dithiothreitol (DTT) to the cell culture medium. This could easily be applied for single-cell studies by fluid exchange with DTT containing medium at defined time points and periods. It would enable studying how the population heterogeneity changes under stress or even for different stress levels and whether some parameters are more affected than others. It is possible, that the populations broadens with increasing misfolding stress, but also a bimodal distribution could develop in order to increase the survival probability of the population.

In conclusion, the presented quantification methods for single-cell protein expression time courses may help to gain a deeper understanding of mRNA delivery processes and the regulation of transfection mediated translation especially the control of mRNA half-life. It was shown that cells can be successfully used as living test tubes to quantify protein expression by delivering exogenous mRNA. The developed single-cell translation assay has the potential to reveal the mechanisms of artificial gene delivery and to gain new insights in degradation regulation, which is both relevant for gene therapy. In regard to the opening quote, I would like to say:

> *"Cells are indeed fascinating factories that produce the hardware of choice if the right instructions are delivered. The question is how much output you are going to get."*

A. Protocols

Cell culture

The liver carcinoma cell line HuH7 (Figure A.1.) was used for the translation studies in the context of this thesis. The cells were cultured at 37 °C in a humidified atmosphere at 5% CO_2 and were passaged after reaching 70-80% confluence. The cells were cultured in the cell growth medium RPMI-1640 containing GlutaMax and supplemented with 10% FBS, 5 mM HEPES, and 1 mM sodium pyruvate.

For passaging, the cells were washed with PBS before adding the detachment solution Accutase. After incubating the detachment solution for 2 min, the cells were resuspended in cell growth medium to detach all cells from the flask's bottom. The cells are collected in a 15 ml reaction tube and centrifuged at room temperature for 3 min at 100 rcf. The cell suspension concentration is meanwhile determined using a Neubauer counting chamber. As the generation time for the cell line was determined to be around 30 h, the cells were seeded in an appropriate density depending on the cultivation times (typically between two to four days).

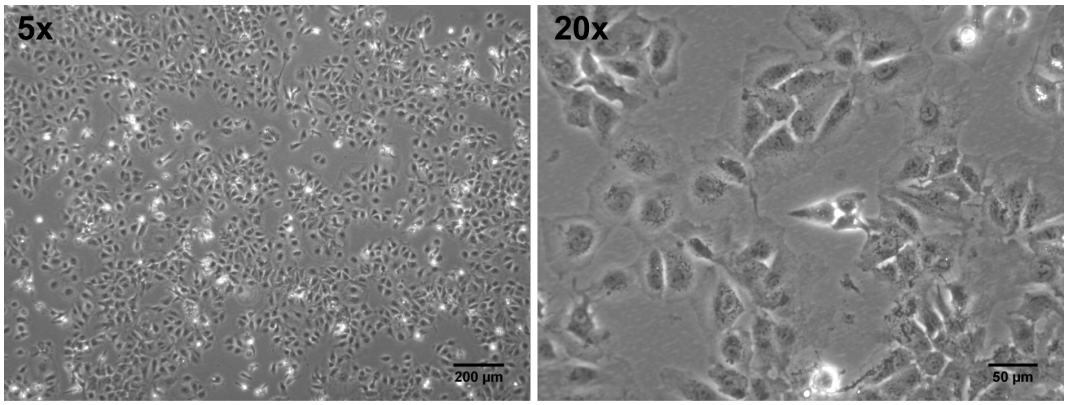

Figure A.1.: The phase contrast images show the liver carcinoma cell line HuH7 at two different magnifications before passaging (5x and 20x respectively).

Single-cell array preparation

Array fabrication

A detailed description of the protocol for the single-cell microarray fabrication using µPIPP is published by Reiser et al. in Cell-based Microarrays: Methods and Protocols

[70]. For all studies using HuH7 cells a pattern geometry with squares of 30 µm edge length and a distance of 60 µm between the squares was used (Figure A.2. A). The slide was prepared the day before the experiment.

Cell seeding

Each washing step mentioned in the protocol consists of adding 120 µl of the denoted liquid to one reservoir of the channel and removing 120 µl on the other side.

1. Remove the PBS in the reservoirs of each channel.
2. Wash each channel with 37 °C warm cell growth medium.
3. Add 60 µl cell growth medium to each channel and put the slide back in incubator.
4. Detach the cells from the culture flask as described in the subsection "Cell culture" and adjust the cell suspension to a concentration of $4{\cdot}10^5$ cells/ml.
5. Add 30 µl cell suspension to each channel and mix the cell growth medium with the cell suspension by removing liquid from one reservoir and adding it to the other to reach a homogeneous cell distribution in the channel (Figure A.2. B).
6. The slide needs to incubate for 4 h in an incubator at 37 °C to guarantee cell adhesion and self-sorting on the micropattern (Figure A.2. C). The single-cell occupancy of the micropattern can be improved by exchanging the cell growth medium 1 h after adding the cell suspension to remove non adherent cells.

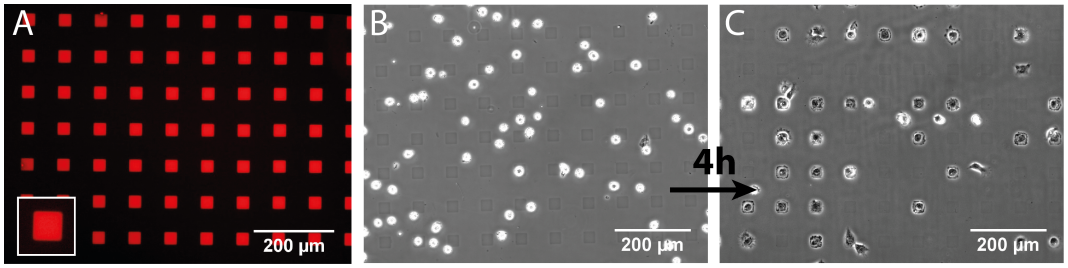

Figure A.2.: Cellular self-organization on µPIPP pattern. (A) Micropattern of FN adhesion squares (red). (B) HuH7 cells are randomly distributed after seeding. (C) 4 h after seeding most cells adhered on the FN squares and aligned to the pattern. Reprinted with permission [70].

Perfusion system

Material

1. Needle-free swabable valves (Mednet)
2. Female luer lock connector (Mednet)
3. Male luer connector (in-house built)
4. PTFE microtubing with an inner diameter of 0.3 mm
5. Silicon tubing with an inner diameter of 0.8 mm

6. Syringe (1 ml with replacement sporn)
7. 6 ml cell culture medium (e.g. L15 supplemented with 10% FBS, without phenol red)
8. 15 ml reaction tube

Inlet — The valve is screwed on a female luer connector. A 2 cm piece of silicon tube is put over a 20 cm piece of PTFE tubing at one end. This end is connected to the female luer. The other end of the PTFE tube is agglutinated with one male luer using a glue that is non toxic for cells (see Figure A.3. A).

Outlet — A PTFE tube of 25 cm length is clued with a male luer at one end (see Figure A.3. B).

Serial — A short PTFE tube piece of 3 cm length is agglutinated to male luers at both ends to obtain a serial connector (see Figure A.3. C).

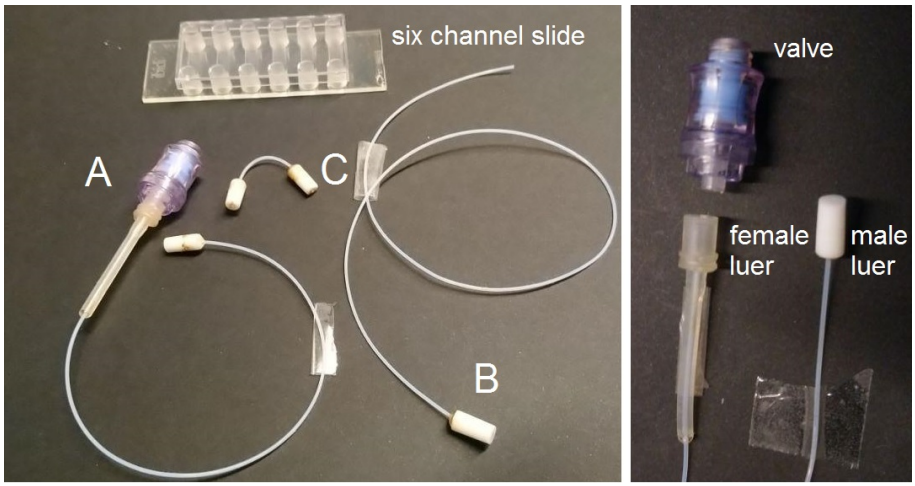

Figure A.3.: The main components of the perfusion system, which is connected to a six channel slide, for liquid handling on microscope. The inlet tube (A) consists of a valve connected to the PTFE tube via a female luer. The other end of the tubing is connected to a male luer. The outlet tube (B) is only connected to one male luer. To connect two channels a short piece of PTFE tube is agglutinated to male luers on both ends (C).

Connection to slide

The connection of the perfusion system has to be done under sterile conditions and performed as fast as possible to avoid cooling of the cells. It is very important to connect all pieces without the inclusion of air bubbles as the bubbles would damage the cells during liquid handling. The cell culture medium used for connection needs to be warmed up to 37 °C to slow down cooling of the cells during connection.

If L15 is used for the time-lapse measurement, remove CO_2 dependent medium by washing each channel once with 120 µl L15. All reservoirs are completely filled with cell culture medium (take care that no air bubble is trapped at the bottom of a reservoir). The inlet tube is completely filled with cell culture medium using the syringe containing 1 ml medium. Connect the inlet tube with the attached syringe to one reservoir without

trapping an air bubble. Next, a serial connector or outlet tube can be connected to the other reservoir of the channel. In both cases a male luer is put into the opposite reservoir to the inlet tube. If an outlet tube was connected, carefully flush medium through the system using the syringe connected to the inlet tube until liquid drips out (check for leaking). If a serial connector was connected fill the short tube and connect the other male luer to the neighboring reservoir. This step can be repeated for more than two channels. Finally, connect an outlet tube as already described to the opposite reservoir. After flushing the outlet tube, the syringe is carefully removed and drawn up to 1 ml to connect the next channel beginning with a new inlet tube until all channels are connected.

Set up on microscope

The slide is put in the warmed up heating box on the microscope stage. To avoid movement of the slide during liquid handling, each tube is fixed with some tape to the stage. Take care that it is still possible to connect a syringe to the valve. The ends of the outlet tubes are inserted through a hole in the cap of a 15 ml reaction tube to collect the liquid waste. For each washing step on stage, a syringe is plugged carefully in one of the swabable valves and the containing liquid is slowly flushed through the perfusion system. Take care that the applied pressure is as low as possible to avoid any damage to the cells or the slide. The use of a syringe with a replacement sporn is recommended for liquids that are only available in small quantities to keep the death volume of the system low.

Reuse of tubing system

The tubing system is reusable if cleaned properly after every use. In order to clean the perfusion system, each piece of the tubing system is rinsed with 70% ethanol and rinsed a second time with distilled water to remove ethanol completely. Thereby, each piece is checked for leaking. In case of leaking the tubing or sealing needs to be replaced. For sterilization, the perfusion system is put in a glass box and autoclaved.

mRNA transfection using Lipofectamine2000

All steps of the protocol for mRNA L2000 transfection are made with nuclease free pipette tips and reaction tubes to avoid mRNA degradation during preparation. The mRNA stock solutions need to be kept on ice during preparation to avoid degradation. The specified volumes in this protocol are for the preparation of 3x 600 µl transfection solution for the on stage transfection during a time-lapse experiment of three tubing systems (each connecting two channels). In general, the ratio of L2000:mRNA of 2.5 µl:1 µg has to be kept constant and the final mRNA concentration during transfection incubation was 0.5 ng/µl.

Reagents

mRNA	Stock solution of IVT mRNA in nuclease free water, stored at -80 °C
L2000	Lipofectamine2000 (Thermo Fisher), stored at 4 °C
OM	Opti-MEM (Thermo Fisher), stored at 4 °C
PBS	pH 7.4, in-house prepared

Lipoplex formation

Prepare the following two solutions:

I. 2.5 µl L2000 stock solution is diluted in a total volume of 500 µl OM, incubation for 5 min at RT

II. 300 ng mRNA per sample in a total volume of 150 µl OM (three samples)

150 µl of solution I. is added to each mRNA dilution (II.) and mixed well by pipetting, resulting in a mRNA-L2000 solution of 1 ng/µl. The mixture incubated for 20 min at RT. After incubation, the mixture is diluted to the final mRNA concentration of 0.5 ng/µl by adding 300 µl OM per mixture.

Tube transfection

The specified volume is for the transfection using one of the tubing systems connected to two channels of the microarray on the microscope. The tubing systems are flushed using a 1 ml syringe with a replacement sporn (to minimize the death volume) per mRNA sample. It is important for each fluid exchange to take care that no air bubble is injected into the tubing system. The air bubble would lead to a damage of the cells on the micropattern, which in the worst case, makes the measurement unusable:

1. Flush the tubing system with 37 °C warm PBS
2. Flush the tubing system with 500 µl of the mRNA-L2000 transfection solution. The transfection solution is incubated for 1 h.
3. Flush out the unbound lipoplexes by washing with 37 °C warm L15 supplemented with 10% FBS

Translation inhibition using cycloheximide

The protocol for inhibiting translation was established previously [118, 119].

Reagents

CHX 15 µg/ml cycloheximide diluted in PBS

L15 Leibowitz 15 (without phenol red)

Protocol

1. A single-cell translation assay has been set up including single-cell array preparation, connection of the perfusion system to the array, and mRNA transfection.
2. For control purposes at least one channel of the array is not treated any further to record unaffected single-cell fluorescence time courses (Figure A.4. A).
3. The other channels are flushed with 300 µl CHX solution per channel 8 h after transfection.
4. The fluorescence intensity courses are recorded up to a total experiment duration of 30 h. The single-cell time course after CHX addition (Figure A.4. B) clearly show different kinetics compared to the control cells.

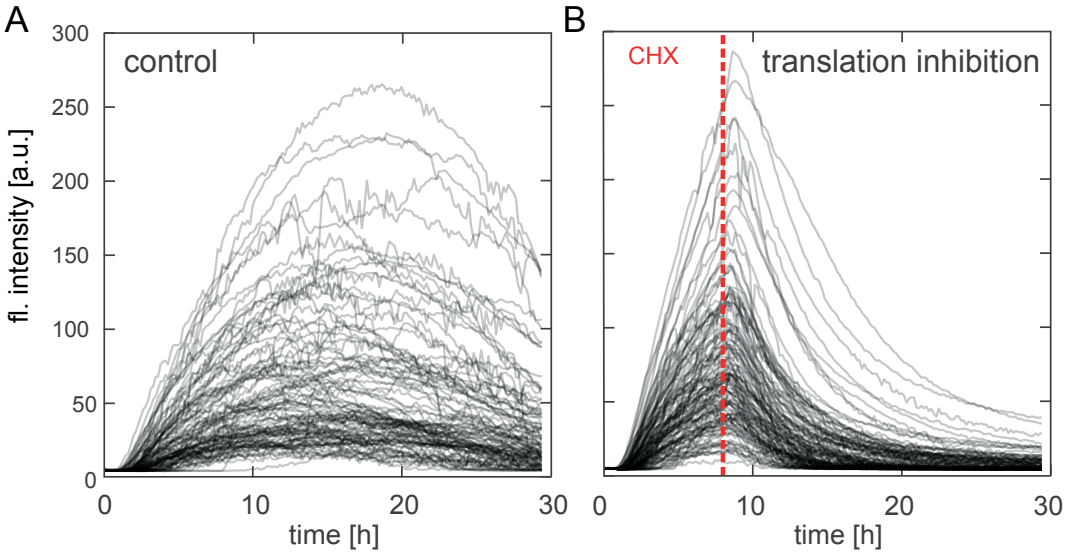

Figure A.4.: Single-cell time courses of cells expressing d2eGFP without translation inhibition (A) and dependent on the addition of CHX (red) 8 h after transfection (B). Adapted from [118].

Fluorescence *in situ* hybridization after transfection

The following protocol is based on the protocol provided by Stellaris for adherent cells [233]. The buffers for washing and hybridization were prepared similarly to a previous study [62].

Reagents

Oligos	Oligonucleotides binding against eGFP sequence labeled with Quasar Dye 670 (Stellaris)
mRNA	eGFP mRNA (Trilink) transfected using L2000
FA	3.7% Formaldehyd solution in PBS
nf H_20	DEPC treated water (0.1% DEPC in ultrapure water, autoclaved); nuclease free (nf)
2x SSC	prepared from 20% SSC stock and diluted in nf water
PBS	pH 7.4, in-house prepared
PBST	PBS supplemented with 0.1% Tween-20
DAPI	nuclei label
Phalloidin	Phalloidin-Rhodamine (Thermo Fisher) for actin labeling

Preparation

FISH probes

The commercial "design ready set (5nmol)" of oligonucleotides labeled with Quasar 670 (Stellaris) to detect mRNA encoding for eGFP was used.

1. Re-dissolve dried oligonucleotides in 400 µl TE buffer (10 mM TRIS, 1 mM EDTA at pH 8.0) to reach a final oligonucleotide concentration of 12.5 µM.
2. Mix well by pipetting the solution up and down, vortex, and centrifuge carefully
3. Prepare aliquots, which can be stored at -20 °C

Wash buffer (10 ml)

Prepare fresh for each experiment and store on ice.

Agent	Volume
Formamide	1.76 ml
20x SSC	1 ml
Tween-20	10 µl
nf water	add up to 10 ml

Hybridization buffer (10 ml)

Buffer is adapted from [62].

Agent	Amount
Dextran sulfat	1 g
Formamide	1.76 ml
20x SSC	1 ml
50 mg/ml BSA	40 µl
200 mM Ribonucleoside vanadyl complex	100 µl
nf water	add up to 10 ml

1. Weigh Dextran in a 50 ml reaction tube
2. Add 5 ml nf water
3. Vortex the tube and put it on a shaker plate for 30-60 min until Dextran is completely dissolved
4. Add other reagents and nf water to reach the final volume of 10 ml
5. Sterile filtrate the buffer and prepare 500 µl aliquots
6. Aliquots can be stored at -20 °C

Fixation

The volumes are given per channel.

1. Wash with 120 µl PBS to remove serum proteins.
2. Incubate with 3.7% FA for 10 min at RT (or 5 min at 37 °C if fixation is done during time-lapse measurement).
3. Wash with 120 µl PBS to remove FA.
4. Incubate cells with a 70% EtOH solution for at least 1 h at 6 °C for permeabilization of the cell membrane.
5. Wash with 2x 120 µl PBS. The slide can be stored at 6 °C until further use.

Hypridization

1. Preparation of oligonucleotide solution: Dissolve the aliquoted stock solution of 12.5 µM to a final concentration of 125 nM in hybridization buffer. A total volume of 250 µl is needed for hybridization of six channels with 40 µl per channel. Protect the solution from light to avoid damage of the fluorescent dyes.
2. Wash with 120 µl PBST.
3. Add 120 µl wash buffer for 5 min.
4. Remove volume in the channels completely.
5. Add 40 µl of the oligonucleotide solutions per channel. Incubate the solution over night at 37 °C. To avoid evaporation put a piece of parafilm between the channel slide and the lid and put the slide in a petri dish with a wet piece of tissue.
6. Add 120 µl wash buffer and incubate in the dark at 37 °C for 30 min.
7. Prepare staining solution containing DAPI and Phalloidin-Rhodamine (Stock with 6.6 µM) in wash buffer with 5 ng/µl DAPI and 150 nM Phalloidin.
8. Add 120 µl staining solution and incubate in the dark at 37 °C for 30 min.
9. Wash twice with 120 µl PBST.
10. The slide is ready for imaging. Alternatively, remove the PBST and fill channels with mounting medium (ibidi).

B. Deterministic translation models

Translation model

The translation model is the least complex model describing the translation and degradation steps as first order kinetics. The mRNA m is translated with rate k_{TL} into GFP G. mRNA and GFP molecules, are degraded by rate δ and β, respectively. The initial amount of mRNA is given by m_0.

$$\begin{aligned}\dot{m} &= -\delta \cdot m \\ \dot{G} &= -\beta \cdot G + k_{TL} \cdot m\end{aligned}$$

Ribosome model

The ribosome model is an extension of the translation model by describing the translation process dependent on ribosome binding R (with the initial concentration R_0), which leads to the formation of a mRNA-ribosome complex mR. The translation thereby dependents on two rates named k_1 and k_2.

$$\begin{aligned}\dot{m} &= -\delta \cdot m - k_1 m \cdot R + k_2 \cdot mR \\ \dot{R} &= -k_1 \cdot m \cdot R + k_2 mR \\ \dot{mR} &= -\dot{R} \\ \dot{G} &= -\beta \cdot G + k_2 \cdot mR\end{aligned}$$

The ribosome kinetics can be approximated as a Michaelis-Menten kinetic describing the enzymatic activity of ribosomes with the parameter α resulting in a simplification of the above ODEs [118]:

$$\begin{aligned}\dot{m} &= -\delta \cdot m \\ \dot{G} &= -\beta \cdot G + k_{TL} \cdot \frac{1}{1 + \alpha \cdot m} \cdot m\end{aligned}$$

Enzyme degradation model

The enzyme degradation model describes the mRNA depletion as an enzymatic step. The degradation enzyme E (initial concentration E_0) binds to the mRNA forming a complex

mE. The degradation kinetic is described by the parameters δ_1 and δ_2. The enzyme degradation model has the same number of parameters as the ribosome model.

$$\begin{aligned}
\dot{m} &= -\delta_1 \cdot m \cdot E \\
\dot{E} &= -\delta_1 \cdot m \cdot E + \delta_2 \cdot mE \\
\dot{mE} &= -\dot{E} \\
\dot{G} &= -\beta \cdot G + k_{TL} \cdot m
\end{aligned}$$

Enzyme degradation-ribosome model

The enzyme degradation-ribosome model is the combination of both models describing the mRNA degradation and translation as an enzymatic reaction instead of first order kinetics.

$$\begin{aligned}
\dot{m} &= -\delta_1 \cdot m \cdot E - k_1 \cdot m \cdot R + k_2 \cdot mR \\
\dot{R} &= -k_1 \cdot m \cdot R + k_2 \cdot mR \\
\dot{mR} &= -\dot{R} \\
\dot{E} &= -\delta_1 \cdot m \cdot E + \delta_2 \cdot mE \\
\dot{mE} &= -\dot{E} \\
\dot{G} &= -\beta \cdot G + k_2 \cdot mR
\end{aligned}$$

Translation-maturation model

The translation-maturation model extends the two step translation model by an additional maturation step. Each of the three steps are modeled as first order kinetics. The mRNA is translated into the unmature variant of GFP G^* that matures into fluorescent GFP with rate k_M.

$$\begin{aligned}
\dot{m} &= -\delta \cdot m \\
\dot{G^*} &= -\beta \cdot G^* - k_M \cdot G^* + k_{TL} \cdot m \\
\dot{G} &= -\beta \cdot G + k_M \cdot G^*
\end{aligned}$$

Abbreviations

ACF	Autocorrelation function
AIC	Akaike Information Criterion
apoE	Apolipoprotein E
BIC	Bayesian Information Criterion
CayRFP	CayenneRFP
Chol	Cholesterol
CHX	Cycloheximide
Cy5	Cyanine dye 5
d2eGFP	Destabilized enhanced green fluorescent protein
DFHBI	3,5-difluoro-HBI
DLin	DLin-DMA
DSPC	Distearoyl phosphocholine
DTT	Dithiothreitol
eGFP	Enhanced green fluorescent protein
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
FISH	Fluorescent in situ hybridization
FL	Fluorescence
FRET	Förster resonance energy transfer
FWHM	Full width half maximum
GFP	Green fluorescent protein
HBI	4-hydroxybenzlidene imidazolinone
IVT	in vitro transcribed
KC2	DLin-KC2-DMA

L2000	Lipofectamine 2000
LISCA	Live cell imaging of single-cell arrays
LNP	Lipid nanoparticle
MC3	DLin-MC3-DMA
MIMIC	Micromolding in capillaries
MLE	Maximum likelihood estimation
mRNA	messenger RNA
MWA	Microwell analysis
NA	Nucleic acid
nf	Nuclease free
NLME	Non-linear mixed effect approach
ODE	Ordinary differential equations
ORF	Open reading frame
PABP	poly(A) binding protein
PDMS	Poly(dimethylsiloxane)
pDNA	Plasmid DNA
PEG	Polyethylenglycol
PEG-DA	PEG-diacrylate
pKa	Logarithmic acid dissociation constant
PLL-g-PEG	Poly-L-lysine grafted poly(ethylene glycol)
RISC	RNA-induced-silencing complex
RT	Room temperature
sem	Standard error of the mean
siRNA	small interfering RNA
std	Standard deviation
STS	Standard two-stage approach
TE	Transfection efficiency
UTR	Untranslated region

References

[1] E. Schrödinger. *Was ist Leben? Die lebende Zelle mit den Augen des Physikers betrachtet*. Piper Verlag, 2010. 1

[2] L. von Bertalanffy. The theory of open systems in physics and biology. *Science*, 111(2872):23–29, 1950. 1

[3] F. Crick. Central dogma of molecular biology. *Nature*, 227(5258):561–563, Aug 1970. 1, 2.1

[4] E. Klipp, R. Herwig, A. Kowald, C. Wierling, and H. Lehrach. *Systems biology in practice*. Wiley-VCH, Weinheim, 2005. 1, 2.2, 4

[5] U. Alon. *An introduction to systems biology: design principles of biological circuits*. Chapman and Hall/CRC, 2006. 1, 2.2

[6] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, Aug 2002. 1, 1, 2.1, 2.2

[7] N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, and M. B. Elowitz. Gene regulation at the single-cell level. *Science*, 307(5717):1962–1965, Mar 2005. 1, 2.2

[8] A. Aderem. Systems biology: its practice and challenges. *Cell*, 121(4):511–513, 2005. 1

[9] A. Sanchez and I. Golding. Genetic determinants and cellular constraints in noisy gene expression. *Science*, 342(6163):1188–1193, Dec 2013. 1

[10] S. O. Olofsson, K. Bostrom, P. Carlsson, J. Boren, M. Wettesten, G. Bjursell, O. Wiklund, and G. Bondjers. Structure and biosynthesis of apolipoprotein B. *Am. Heart J.*, 113(2 Pt 2):446–452, Feb 1987. 1

[11] K. Lundstrom. Latest development on RNA-based drugs and vaccines. *Future Sci OA*, 4(5):FSO300, Jun 2018. 1

[12] H. Ledford. Gene-silencing technology gets first drug approval after 20-year wait. *Nature*, 560(7718):291–292, Aug 2018. 1, 5.1, 7.1

[13] K. Servick. Runnders up for breakthrough of the year: Gene-silencing drug approved. https://vis.sciencemag.org/breakthrough2018/finalists/#rna-drug, 2018. 1

[14] U. Sahin, K. Kariko, and O. Tureci. mRNA-based therapeutics–developing a new class of drugs. *Nat Rev Drug Discov*, 13(10):759–780, Oct 2014. 1, 4.5, 5

[15] K. A. Hajj and K. A. Whitehead. Tools for translation: non-viral materials for therapeutic mRNA delivery. *Nat. Rev. Mat.*, 2(10):17056, 2017. 1, 2.1, 5.1, 5.1, 5.1, 5.1

[16] L. M. P. Vermeulen, T. Brans, S. C. De Smedt, K. Remaut, and K. Braeckmans. Methodologies to investigate intracellular barriers for nucleic acid delivery in non-viral gene therapy. *Nano Today*, 21:74 – 90, 2018. 1, 5

[17] D. Y. Kim, J. S. Kwon, J. H. Lee, L. M. Jin, J. H. Kim, and M. S. Kim. Effects of the Surface Charge of Stem Cell Membranes and DNA/Polyethyleneimine Nanocomplexes on Gene Transfection Efficiency. *J Biomed Nanotechnol*, 11(3):522–530, Mar 2015. 1, 5.1

[18] K. J. Kauffman, M. J. Webber, and D. G. Anderson. Materials for non-viral intracellular delivery of messenger RNA therapeutics. *J Control Release*, 240:227–234, 10 2016. 1, 5

[19] W. Sung. Introduction: Biological systems and physical approaches. In *Statistical Physics for Biological Matter*, pages 1–6. Springer, 2018. 1

[20] T. S. Ligon, C. Leonhardt, and J. O. Rädler. Multi-level kinetic model of mRNA delivery via transfection of lipoplexes. *PLoS ONE*, 9(9):e107148, 2014. 1, 2.2, 4.1

[21] C. Leonhardt, G. Schwake, T. R. Stogbauer, S. Rappl, J. T. Kuhr, T. S. Ligon, and J. O. Rädler. Single-cell mRNA transfection studies: delivery, kinetics and statistics by numbers. *Nanomedicine*, 10(4):679–688, May 2014. 1, 2.2, 2.3, 2.4, 3, 4, 4.1, 4.1, 5.2

[22] S. J. Altschuler and L. F. Wu. Cellular heterogeneity: do differences make a difference? *Cell*, 141(4):559–563, May 2010. 1, 2.2

[23] D. Di Carlo, L. Y. Wu, and L. P. Lee. Dynamic single cell culture array. *Lab Chip*, 6(11):1445–1449, Nov 2006. 1

[24] J. M. Levsky and R. H. Singer. Gene expression and the myth of the average cell. *Trends Cell Biol.*, 13(1):4–6, Jan 2003. 1

[25] X. S. Xie, J. Yu, and W. Y. Yang. Living cells as test tubes. *Science*, 312(5771):228–230, Apr 2006. 1

[26] E. Pennisi. Breakthrough of the year: Development cell by cell. https://vis.sciencemag.org/breakthrough2018/finalists/, 2018. 1

[27] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach. Global quantification of mammalian gene expression control. *Nature*, 473(7347):337–342, May 2011. 1, 2.1, 2.2

[28] C. Vogel, R. d. e. S. Abreu, D. Ko, S. Y. Le, B. A. Shapiro, S. C. Burns, D. Sandhu, D. R. Boutz, E. M. Marcotte, and L. O. Penalva. Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. *Mol. Syst. Biol.*, 6:400, Aug 2010. 1, 2.1

[29] Z. Cheng, G. Teo, S. Krueger, T. M. Rock, H. W. Koh, H. Choi, and C. Vogel. Differential dynamics of the mammalian mRNA and protein expression response to misfolding stress. *Mol. Syst. Biol.*, 12(1):855, Jan 2016. 1, 8

[30] H. C. Fan, G. K. Fu, and S. P. Fodor. Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. *Science*, 347(6222):1258367, Feb 2015. 1

[31] D. Muzzey and A. van Oudenaarden. Quantitative time-lapse fluorescence microscopy in single cells. *Annu. Rev. Cell Dev. Biol.*, 25:301–327, 2009. 1, 2.2

[32] J. C. Locke and M. B. Elowitz. Using movies to analyse gene circuit dynamics in single cells. *Nat. Rev. Microbiol.*, 7(5):383–392, May 2009. 1, 2.2, 3

[33] M. Ferizi, C. Leonhardt, C. Meggle, M. K. Aneja, C. Rudolph, C. Plank, and J. O. Rädler. Stability analysis of chemically modified mRNA using micropattern-based single-cell arrays. *Lab Chip*, 15(17):3561–3571, Sep 2015. 1, 2.2, 2.3, 3, 3.1.1, 4, 4.1, 5.2, 6

[34] S. Lindstrom and H. Andersson-Svahn. Overview of single-cell analyses: microdevices and applications. *Lab Chip*, 10(24):3363–3372, Dec 2010. 1

[35] J. R. Rettig and A. Folch. Large-scale single-cell trapping and imaging using microwell arrays. *Anal. Chem.*, 77(17):5628–5634, Sep 2005. 1, 3.1.3

[36] M. Kaiser, F. Jug, T. Julou, S. Deshpande, T. Pfohl, O. K. Silander, G. Myers, and E. van Nimwegen. Monitoring single-cell gene regulation under dynamically controllable conditions with integrated microfluidics and software. *Nat Commun*, 9(1):212, 01 2018. 1

[37] C. Leonhardt. *On Quantitative mRNA Transfection.* PhD thesis, Ludwig-Maximilians-University Munich, Jun 2014. 1, 2.3, 3.2.2

[38] A. L. Jalkanen, S. J. Coleman, and J. Wilusz. Determinants and implications of mRNA poly(A) tail size–does this protein make my tail look big? *Semin. Cell Dev. Biol.*, 34:24–32, Oct 2014. 1, 6, 6.1.2, 6.3.2

[39] C. C. Friedel, L. Dolken, Z. Ruzsics, U. H. Koszinowski, and R. Zimmer. Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. *Nucleic Acids Res.*, 37(17):e115, Sep 2009. 2.1

[40] J. J. Li, P. J. Bickel, and M. D. Biggin. System wide analyses have underestimated protein abundances and the importance of transcription in mammals. *PeerJ*, 2:e270, 2014. 2.1

[41] J. J. Li and M. D. Biggin. Gene expression. Statistics requantitates the central dogma. *Science*, 347(6226):1066–1067, Mar 2015. 2.1

[42] J. B. Plotkin. Transcriptional regulation is only half the story. *Mol. Syst. Biol.*, 6:406, Aug 2010. 2.1, 2.1

[43] C. Vogel. Translation's coming of age. *Mol. Syst. Biol.*, 7:498, May 2011. 2.1, 2.1, 6

[44] R. de Sousa Abreu, L. O. Penalva, E. M. Marcotte, and C. Vogel. Global signatures of protein and mRNA expression levels. *Mol Biosyst*, 5(12):1512–1526, Dec 2009. 2.1

[45] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.*, 6(6):451–464, Jun 2005. 2.1

[46] G. Tavernier, O. Andries, J. Demeester, N. N. Sanders, S. C. De Smedt, and J. Rejman. mRNA as gene therapeutic: how to control protein expression. *J Control Release*, 150(3):238–247, Mar 2011. 2.1

[47] H. Youn and J. K. Chung. Modified mRNA as an alternative to plasmid DNA (pDNA) for transcript replacement and vaccination therapy. *Expert Opin Biol Ther*, 15(9):1337–1348, 2015. 2.1, 6

[48] C. E. Thomas, A. Ehrhardt, and M. A. Kay. Progress and problems with the use of viral vectors for gene therapy. *Nat. Rev. Genet.*, 4(5):346–358, May 2003. 2.1

[49] O. Shimomura and O. Shimomura. Discovery of green fluorescent protein (GFP) (Nobel Lecture). *Angew. Chem. Int. Ed. Engl.*, 48(31):5590–5602, 2009. 2.1

[50] J. J. van Thor. Photoreactions and dynamics of the green fluorescent protein. *Chem Soc Rev*, 38(10):2935–2950, Oct 2009. 2.1

[51] R. Iizuka, M. Yamagishi-Shirasaki, and T. Funatsu. Kinetic study of de novo chromophore maturation of fluorescent proteins. *Anal. Biochem.*, 414(2):173–178, Jul 2011. 2.1, 5.3.1

[52] J. S. Paige, K. Y. Wu, and S. R. Jaffrey. RNA mimics of green fluorescent protein. *Science*, 333(6042):642–646, Jul 2011. 2.1, 7.3

[53] M. Zimmer. Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. *Chem. Rev.*, 102(3):759–781, Mar 2002. 2.1

[54] F. Buettner, K. N. Natarajan, F. P. Casale, V. Proserpio, A. Scialdone, F. J. Theis, S. A. Teichmann, J. C. Marioni, and O. Stegle. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol.*, 33(2):155–160, Feb 2015. 2.2

[55] B. Snijder and L. Pelkmans. Origins of regulated cell-to-cell variability. *Nat. Rev. Mol. Cell Biol.*, 12(2):119–125, 02 2011. 2.2, 4.5

[56] M. Etzrodt, M. Endele, and T. Schroeder. Quantitative single-cell approaches to stem cell research. *Cell Stem Cell*, 15(5):546–558, Nov 2014. 2.2

[57] P. Dalerba, T. Kalisky, D. Sahoo, P. S. Rajendran, M. E. Rothenberg, A. A. Leyrat, S. Sim, J. Okamoto, D. M. Johnston, D. Qian, M. Zabala, J. Bueno, N. F. Neff, J. Wang, A. A. Shelton, B. Visser, S. Hisamori, Y. Shimono, M. van de Wetering, H. Clevers, M. F. Clarke, and S. R. Quake. Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat. Biotechnol.*, 29(12):1120–1127, Nov 2011. 2.2

[58] B. Bodenmiller, E. R. Zunder, R. Finck, T. J. Chen, E. S. Savig, R. V. Bruggner, E. F. Simonds, S. C. Bendall, K. Sachs, P. O. Krutzik, and G. P. Nolan. Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nat. Biotechnol.*, 30(9):858–867, Sep 2012. 2.2

[59] J. Uhlendorf, A. Miermont, T. Delaveau, G. Charvin, F. Fages, S. Bottani, G. Batt, and P. Hersen. Long-term model predictive control of gene expression at the population and single-cell levels. *Proc. Natl. Acad. Sci. U.S.A.*, 109(35):14271–14276, Aug 2012. 2.2

[60] S. L. Spencer, S. Gaudet, J. G. Albeck, J. M. Burke, and P. K. Sorger. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature*, 459(7245):428–432, May 2009. 2.2

[61] N. Navin, J. Kendall, J. Troge, P. Andrews, L. Rodgers, J. McIndoo, K. Cook, A. Stepansky, D. Levy, D. Esposito, L. Muthuswamy, A. Krasnitz, W. R. McCombie, J. Hicks, and M. Wigler. Tumour evolution inferred by single-cell sequencing. *Nature*, 472(7341):90–94, Apr 2011. 2.2

[62] S. O. Skinner, H. Xu, S. Nagarkar-Jaiswal, P. R. Freire, T. P. Zwaka, and I. Golding. Single-cell analysis of transcription kinetics across the cell cycle. *Elife*, 5:e12175, Jan 2016. 2.2, 7.3.2, A, A

[63] E. I. Chatzopoulou, P. Raharja-Liu, A. Murschhauser, F. Sekhavati, F. Buggenthin, A. M. Vollmar, C. Marr, and J. O. Rädler. A single-cell micro-trench platform for automatic monitoring of cell division and apoptosis after chemotherapeutic drug administration. *Sci Rep*, 8(1):18042, Dec 2018. 2.2

[64] H. Kitano. Systems biology: A brief overview. *Science*, 295(5560):1662–1664, Mar. 2002. 2.2

[65] D. Di Carlo and L. P. Lee. Dynamic single-cell analysis for quantitative biology. *Anal. Chem.*, 78(23):7918–7925, Dec 2006. 2.2, 3.3

[66] A. Raue, M. Schilling, J. Bachmann, A. Matteson, M. Schelke, D. Kaschek, S. Hug, C. Kreutz, B. D. Harms, F. J. Theis, and et al. Lessons learned from quantitative dynamical modeling in systems biology. *PLoS ONE*, 8(9):e74335, Sep 2013. 2.2

[67] G. Schwake, S. Youssef, J. T. Kuhr, S. Gude, M. P. David, E. Mendoza, E. Frey, and J. O. Rädler. Predictive modeling of non-viral gene transfer. *Biotechnol. Bioeng.*, 105(4):805–813, Mar 2010. 2.2, 7.1, 7.1

[68] F. Fröhlich. *Scalable Simulation and Optimization Methods for Differential Equation Models describing Biochemical Reaction Networks*. PhD thesis, Technical University Munich, Apr 2018. 2.2

[69] P. J. Röttgermann, A. P. Alberola, and J. O. Rädler. Cellular self-organization on micro-structured surfaces. *Soft Matter*, 10(14):2397–2404, Apr 2014. 2.3, 3.1.1

[70] A. Reiser, M. L. Zorn, A. Murschhauser, and J. O. Rädler. *Single Cell Microarrays Fabricated by Microscale Plasma-Initiated Protein Patterning (muPIPP)*, chapter 4, pages 41–54. Springer New York, New York, NY, 2018. 2.3, 2.3.1, 3.1, A, A.2

[71] M. Théry. Micropatterning as a tool to decipher cell morphogenesis and functions. *J. Cell. Sci.*, 123(Pt 24):4201–4213, Dec 2010. 2.3.1, 3, 3.1

[72] P. Ertl and M. Rothbauer, editors. *Cell-Based Microarrays*, volume 1771 of *Methods in Molecular Biology*. Humana Press, 2018. 2.3.1, 3.1

[73] C. Schreiber, F. J. Segerer, E. Wagner, A. Roidl, and J. O. Rädler. Ring-Shaped Microlanes and Chemical Barriers as a Platform for Probing Single-Cell Migration. *Sci Rep*, 6:26858, 2016. 2.3.1, 3, 3.3

[74] E. I. Chatzopoulou, C. C. Roskopf, F. Sekhavati, T. A. Braciak, N. C. Fenn, K. P. Hopfner, F. S. Oduncu, G. H. Fey, and J. O. Rädler. Chip-based platform for dynamic analysis of NK cell cytolysis mediated by a triplebody. *Analyst*, 141(7):2284–2295, Apr 2016. 2.3.1, 3.1.1

[75] P. J. Röttgermann, K. A. Dawson, and J. O. Rädler. Time-resolved study of nanoparticle induced apoptosis using microfabricated single cell arrays. *Microarrays*, 5(2):8, 2016. 2.3.1, 3, 3.1.1, 5.6, 7

[76] M. Han, Y. Bae, N. Nishiyama, K. Miyata, M. Oba, and K. Kataoka. Transfection study using multicellular tumor spheroids for screening non-viral polymeric gene

vectors with low cytotoxicity and high transfection efficiencies. *J Control Release*, 121(1-2):38–48, Aug 2007. 2.3.1

[77] T. Kato, M. Tanaka, and M. Oba. Protein transfection study using multicellular tumor spheroids of human hepatoma Huh-7 cells. *PLoS ONE*, 8(12):e82876, 2013. 2.3.1

[78] M. Schwarzfischer, C. Marr, J. Krumsiek, P. S. Hoppe, T. Schroeder, and F. J. Theis. Efficient fluorescence image normalization for time lapse movies. pages http://www.miaab.org/miaab–2011–. MIAAB, 2011. 2.3.4

[79] L. Lechtenberg. Model selection in deterministic models of mRNA transfection. Master's thesis, Ludwig-Maximilians-University Munich, April 2015. 2.3.4, 2.4, 2.4

[80] D. Woschée. The dynamics of apoptosis on the single-cell basis. Master's thesis, Ludwig-Maximilians-University Munich, June 2016. 2.3.4, 2.4

[81] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007. 2.4

[82] J. Vanlier, C.A. Tiemann, P.A.J. Hilbers, and N.A.W. van Riel. Parameter uncertainty in biochemical models described by ordinary differential equations. *Mathematical Biosciences*, 246(2):305–314, Dec 2013. 2.4

[83] F. Fröhlich, A. Reiser, L. Fink, D. Woschée, T. Ligon, F. J. Theis, J. O. Rädler, and J. Hasenauer. Multi-experiment nonlinear mixed effect modeling of single-cell translation kinetics after transfection. *NPJ Syst Biol Appl*, 5:1, Dez 2018. 2.4, 2.4, 3, 4, 5.2

[84] I. J. Myung. Tutorial on maximum likelihood estimation. *J. Math. Psychol.*, 47(1):90–100, Feb 2003. 2.4

[85] A. Raj and A. van Oudenaarden. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*, 135(2):216–226, Oct 2008. 3

[86] F. J. Segerer, F. Thuroff, A. Piera Alberola, E. Frey, and J. O. Rädler. Emergence and Persistence of Collective Cell Migration on Small Circular Micropatterns. *Phys. Rev. Lett.*, 114(22):228102, Jun 2015. 3, 3.1.1

[87] S. Gruber. Rauschanalyse von Fluoreszenzeinzelzellspuren auf Mikrostrukturen. Bachelor's thesis, May 2016. 3

[88] M. Piel and M. Théry, editors. *Micropatterning in Cell Biology*, volume 119. Academic Press, 2014. 3.1

[89] X. Jiang, D. A. Bruzewicz, A. P. Wong, M. Piel, and G. M. Whitesides. Directing cell migration with asymmetric micropatterns. *Proc. Natl. Acad. Sci. U.S.A.*, 102(4):975–978, Jan 2005. 3.1

[90] Anais Pitto-Barry and Nicolas P. E. Barry. Pluronic block-copolymers in medicine: from chemical and biological versatility to rationalisation and clinical advances. *Polym. Chem.*, 5:3291–3297, 2014. 3.1.2

[91] J. Stephan, F. Keber, V. Stierle, J. O. Rädler, and P. Paulitschke. Single-Cell Optical Distortion Correction and Label-Free 3D Cell Shape Reconstruction on Lattices of Nanostructures. *Nano Lett.*, 17(12):8018–8023, 12 2017. 3.1.2

[92] E. Kim and G. M. Xia, Y.and Whitesides. Polymer microstructures formed by moulding in capillaries. *Nature*, 376(6541):581, 1995. 3.1.3

[93] M. Ochsner, M. R. Dusseiller, H. M. Grandin, S. Luna-Morris, M. Textor, V. Vogel, and M. L. Smith. Micro-well arrays for 3D shape control and high resolution analysis of single cells. *Lab Chip*, 7(8):1074–1077, Aug 2007. 3.1.3

[94] F. Sekhavati, M. Endele, S. Rappl, A. K. Marel, T. Schroeder, and J. O. Rädler. Marker-free detection of progenitor cell differentiation by analysis of Brownian motion in micro-wells. *Integr Biol (Camb)*, 7(2):178–183, Feb 2015. 3.1.3

[95] F. Sekhavati. *Dynamic response of individual cells in heterogeneous population*. PhD thesis, Ludwig-Maximilians-University Munich, May 2015. 3.2

[96] E. Chatzopoulo. *Single-cell arrays for dynamic analysis of cell processes*. PhD thesis, Ludwig-Maximilians-University Munich, Jun 2018. 3.1.3, 3.2.1

[97] I. Begemann, T. Saha, L. Lamparter, I. Rathmann, D. Grill, L. Golbach, C. Rasch, U. Keller, B. Trappmann, M. Matis, V. Gerke, J. Klingauf, and M. Galic. Mechanochemical self-organization determines search pattern in migratory cells. *Nature Physics*, page 1, 2019. 3.3

[98] P. J. Röttgermann, S. Hertrich, I. Berts, M. Albert, F. J. Segerer, J. F. Moulin, B. Nickel, and J. O. Rädler. Cell motility on polyethylene glycol block copolymers correlates to fibronectin surface adsorption. *Macromol Biosci*, 14(12):1755–1763, Dec 2014. 3.3

[99] P. Corish and C. Tyler-Smith. Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Eng.*, 12(12):1035–1040, Dec 1999. 4, 4.1, 4.3, 4.5, 6.2.2

[100] J. Almquist, L. Bendrioua, C. B. Adiels, M. Goksor, S. Hohmann, and M. Jirstrand. A Nonlinear Mixed Effects Approach for Modeling the Cell-To-Cell Variability of Mig1 Dynamics in Yeast. *PLoS ONE*, 10(4):e0124050, 2015. 4, 4.2.1, 4.5

[101] M. Karlsson, D. L. Janzen, L. Durrieu, A. Colman-Lerner, M. C. Kjellsson, and G. Cedersund. Nonlinear mixed-effects modelling for single cell estimation: when, why, and how to use it. *BMC Syst Biol*, 9:52, Sep 2015. 4, 4.2.1, 4.5

[102] M. K. Kalita, K. Sargsyan, B. Tian, A. Paulucci-Holthauzen, H. N. Najm, B. J. Debusschere, and A. R. Brasier. Sources of cell-to-cell variability in canonical nuclear factor-Î¹B (NF-Î¹B) signaling pathway inferred from single cell dynamic images. *J. Biol. Chem.*, 286(43):37741–37757, Oct 2011. 4

[103] L. B. Sheiner and S. L. Beal. Evaluation of methods for estimating population pharmacokinetic parameters. III. Monoexponential model: routine clinical pharmacokinetic data. *J Pharmacokinet Biopharm*, 11(3):303–319, Jun 1983. 4

[104] J. C. Pinheiro. *Topics in Mixed Effects Models*. PhD thesis, University of Wisconsin - Madison, 1994. 4

[105] C. Zechner, M. Unger, S. Pelet, M. Peter, and H. Koeppl. Scalable inference of heterogeneous reaction kinetics from pooled single-cell recordings. *Nat. Methods*, 11:197–202, Jan. 2014. 4, 4.4, 4.5

[106] A. Llamosi, A. M. Gonzalez-Vargas, C. Versari, E. Cinquemani, G. Ferrari-Trecate, P. Hersen, and G. Batt. What Population Reveals about Individual Cell Identity:

Single-Cell Parameter Estimation of Models of Gene Expression in Yeast. *PLoS Comput. Biol.*, 12(2):e1004706, Feb 2016. 4, 4.5

[107] M. S. Kormann, G. Hasenpusch, M. K. Aneja, G. Nica, A. W. Flemmer, S. Herber-Jonat, M. Huppmann, L. E. Mays, M. Illenyi, A. Schams, M. Griese, I. Bittmann, R. Handgretinger, D. Hartl, J. Rosenecker, and C. Rudolph. Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. *Nat. Biotechnol.*, 29(2):154–157, Feb 2011. 4.1, 6.1.3

[108] A. Wahlers, M. Schwieger, Z. Li, D. Meier-Tackmann, C. Lindemann, H. G. Eckert, D. von Laer, and C. Baum. Influence of multiplicity of infection and protein stability on retroviral vector-mediated gene expression in hematopoietic cells. *Gene Ther.*, 8(6):477–486, Mar 2001. 4.1, 4.3

[109] Y. Liu, A. Beyer, and R. Aebersold. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell*, 165(3):535–550, Apr 2016. 4.1

[110] O.-T. Chis, J. R. Banga, and E. Balsa-Canto. Structural identifiability of systems biology models: A critical comparison of methods. *PLoS ONE*, 6(11):e27755, Nov. 2011. 4.2.1

[111] S. Darmanis, C. J. Gallant, V. D. Marinescu, M. Niklasson, A. Segerman, G. Flamourakis, S. Fredriksson, E. Assarsson, M. Lundberg, S. Nelander, and U. Westermark, B.and Landegren. Simultaneous multiplexed measurement of RNA and proteins in single cells. *Cell Reports*, 14(2):380–389, 01 2016. 4.2.1

[112] A. P. Frei, F. A. Bava, E. R. Zunder, E. W. Hsieh, S. Y. Chen, G. P. Nolan, and P. F. Gherardini. Highly multiplexed simultaneous detection of RNAs and proteins in single cells. *Nat. Methods*, 13(3):269–275, Mar 2016. 4.2.1

[113] A. Raue, C. Kreutz, F. J. Theis, and J. Timmer. Joining forces of Bayesian and frequentist methodology: A study for inference in the presence of non-identifiability. *Philos T Roy Soc A*, 371(1984), Jan. 2013. 4.2.1

[114] T. S. Ligon, F. Fröhlich, O. T. Chi, J. R. Banga, E. Balsa-Canto, and J. Hasenauer. Genssi 2.0: Multi-experiment structural identifiability analysis of sbml models. *Bioinformatics*, page btx735, 2017. 4.2.1

[115] M. Sari, I. Bahceci, U. Dogrusoz, S. O. Sumer, B. A. Aksoy, O. Babur, and E. Demir. SBGNViz: A Tool for Visualization and Complexity Management of SBGN Process Description Maps. *PLoS ONE*, 10(6):e0128985, 2015. 4.5

[116] R. Parker and H. Song. The enzymes and control of eukaryotic mrna turnover. *Nat. Struct. Mol. Biol.*, 11:121 EP –, 02 2004. 4.3

[117] X. Li, X. Zhao, Y. Fang, X. Jiang, T. Duong, C. Fan, C. C. Huang, and S. R. Kain. Generation of destabilized green fluorescent protein as a transcription reporter. *J. Biol. Chem.*, 273(52):34970–34975, Dec 1998. 4.3, 6.2.2

[118] S. Ezendam. Quantification of mRNA stability depending on poly(A) tail length. Master's thesis, Ludwig-Maximilians-Universtity Munich, Dez 2018. 4.3, 6, 6.2.2, A, A.4, B

[119] R. S. Krzyszton. *Towards efficient siRNA delivery and gene silencing kinetics on*

the single cell level. PhD thesis, Ludwig-Maximilians-University, 2018. 4.3, 5.3.1, 7.1, 7.1, A

[120] H. Akaike. *Information Theory and an Extension of the Maximum Likelihood Principle*, pages 199–213. Springer New York, New York, NY, 1998. 4.3

[121] G. Schwarz. Estimating the dimension of a model. *Ann. Statist.*, 6(2):461–464, 1978. 4.3

[122] A. Singh. Transient changes in intercellular protein variability identify sources of noise in gene expression. *Biophys. J.*, 107(9):2214–2220, Nov 2014. 4.5

[123] A. Yamamoto, M. Kormann, J. Rosenecker, and C. Rudolph. Current prospects for mRNA gene delivery. *Eur J Pharm Biopharm*, 71(3):484–489, Mar 2009. 4.5

[124] S. Kreiter, M. Diken, A. Selmi, O. Tureci, and U. Sahin. Tumor vaccination using messenger RNA: prospects of a future therapy. *Curr. Opin. Immunol.*, 23(3):399–406, Jun 2011. 4.5

[125] H. Yin, R. L. Kanasty, A. A. Eltoukhy, A. J. Vegas, J. R. Dorkin, and D. G. Anderson. Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.*, 15(8):541–555, Aug 2014. 5

[126] A. Wittrup and J. Lieberman. Knocking down disease: a progress report on siRNA therapeutics. *Nat. Rev. Genet.*, 16(9):543–552, Sep 2015. 5

[127] W. Zhao, C. Zhang, B. Li, X. Zhang, X. Luo, C. Zeng, W. Li, M. Gao, and Y. Dong. Lipid Polymer Hybrid Nanomaterials for mRNA Delivery. *Cell Mol Bioeng*, 11(5):397–406, Oct 2018. 5

[128] K. A. Whitehead, R. Langer, and D. G. Anderson. Knocking down barriers: advances in siRNA delivery. *Nat Rev Drug Discov*, 8(2):129–138, Feb 2009. 5, 7.1

[129] M. A. Islam, E. K. Reesor, Y. Xu, H. R. Zope, B. R. Zetter, and J. Shi. Biomaterials for mRNA delivery. *Biomater Sci*, 3(12):1519–1533, Dec 2015. 5

[130] J. C. Kaczmarek, P. S. Kowalski, and D. G. Anderson. Advances in the delivery of RNA therapeutics: from concept to clinical reality. *Genome Med*, 9(1):60, 06 2017. 5

[131] G. Sahay, W. Querbes, C. Alabi, A. Eltoukhy, S. Sarkar, C. Zurenko, E. Karagiannis, K. Love, D. Chen, R. Zoncu, Y. Buganim, A. Schroeder, R. Langer, and D. G. Anderson. Efficiency of siRNA delivery by lipid nanoparticles is limited by endocytic recycling. *Nat. Biotechnol.*, 31(7):653–658, Jul 2013. 5

[132] A. J. Lin, N. L. Slack, A. Ahmad, C. X. George, C. E. Samuel, and C. R. Safinya. Three-dimensional imaging of lipid gene-carriers: membrane charge density controls universal transfection behavior in lamellar cationic liposome-DNA complexes. *Biophys. J.*, 84(5):3307–3316, May 2003. 5

[133] R. Bausinger, K. von Gersdorff, K. Braeckmans, M. Ogris, E. Wagner, C. Brauchle, and A. Zumbusch. The transport of nanosized gene carriers unraveled by live-cell imaging. *Angew. Chem. Int. Ed. Engl.*, 45(10):1568–1572, Feb 2006. 5

[134] J. Gilleron, W. Querbes, A. Zeigerer, A. Borodovsky, G. Marsico, U. Schubert, K. Manygoats, S. Seifert, C. Andree, M. Stoter, H. Epstein-Barash, L. Zhang, V. Koteliansky, K. Fitzgerald, E. Fava, M. Bickle, Y. Kalaidzidis, A. Akinc,

M. Maier, and M. Zerial. Image-based analysis of lipid nanoparticle-mediated siRNA delivery, intracellular trafficking and endosomal escape. *Nat. Biotechnol.*, 31(7):638–646, Jul 2013. 5, 5.1

[135] A. Wittrup, A. Ai, X. Liu, P. Hamar, R. Trifonova, K. Charisse, M. Manoharan, T. Kirchhausen, and J. Lieberman. Visualizing lipid-formulated siRNA release from endosomes and target gene knockdown. *Nat. Biotechnol.*, 33(8):870–876, Aug 2015. 5, 5.6, 7.2, 7.4

[136] J. L. Kirschman, S. Bhosle, D. Vanover, E. L. Blanchard, K. H. Loomis, C. Zurla, K. Murray, B. C. Lam, and P. J. Santangelo. Characterizing exogenous mRNA delivery, trafficking, cytoplasmic release and RNA-protein correlations at the level of single cells. *Nucleic Acids Res.*, 45(12):e113, Jul 2017. 5, 7.3.2, 7.4

[137] Y. Wang and L. Huang. A window onto siRNA delivery. *Nat. Biotechnol.*, 31(7):611–612, Jul 2013. 5, 5.6

[138] M. J. W. Evers, J. A. Kulkarni, R. van der Meel, P. R. Cullis, P. Vader, and R. M. Schiffelers. State-of-the-art design and rapid-mixing production techniques of lipid nanoparticles for nucleic acid delivery. *Small Methods*, 2(9):1700375, 2018. 5, 5.1, 5.6

[139] C. Wan, T. M. Allen, and P. R. Cullis. Lipid nanoparticle delivery systems for siRNA-based therapeutics. *Drug Deliv Transl Res*, 4(1):74–83, Feb 2014. 5, 5.1

[140] A. Reiser, D. Woschée, N. Mehrotra, R. Krzyszton, H. H. Strey, and J. O. Rädler. Correlation of mRNA delivery timing and protein expression in lipid-based transfection. *bioRxiv*, 2019. 5, 7.3.1

[141] T. M. Allen and P. R. Cullis. Liposomal drug delivery systems: from concept to clinical applications. *Adv. Drug Deliv. Rev.*, 65(1):36–48, Jan 2013. 5.1, 5.1

[142] H. I. Chang and M. K. Yeh. Clinical development of liposome-based drugs: formulation, characterization, and therapeutic efficacy. *Int J Nanomedicine*, 7:49–60, 2012. 5.1

[143] G. J. Dimitriadis. Translation of rabbit globin mRNA introduced by liposomes into mouse lymphocytes. *Nature*, 274(5674):923–924, Aug 1978. 5.1

[144] D. Pozzi, C. Marchini, F. Cardarelli, F. Salomone, S. Coppola, M. Montani, M. E. Zabaleta, M. A. Digman, E. Gratton, V. Colapicchioni, and G. Caracciolo. Mechanistic evaluation of the transfection barriers involved in lipid-mediated gene delivery: interplay between nanostructure and composition. *Biochim. Biophys. Acta*, 1838(3):957–967, Mar 2014. 5.1, 5.1, 5.6

[145] I. Koltover, T. Salditt, J. O. Rädler, and C. R. Safinya. An inverted hexagonal phase of cationic liposome-DNA complexes related to DNA release and delivery. *Science*, 281(5373):78–81, Jul 1998. 5.1

[146] C. Marchini, D. Pozzi, M. Montani, C. Alfonsi, A. Amici, H. Amenitsch, S. Candeloro De Sanctis, and G. Caracciolo. Tailoring lipoplex composition to the lipid composition of plasma membrane: a Trojan horse for cell entry? *Langmuir*, 26(17):13867–13873, Sep 2010. 5.1

[147] B. Dalby, S. Cates, A. Harris, E. C. Ohki, M. L. Tilkins, P. J. Price, and V. C. Cic-

carone. Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications. *Methods*, 33(2):95–103, Jun 2004. 5.1

[148] A. Masotti, G. Mossa, C. Cametti, G. Ortaggi, A. Bianco, N. D. Grosso, D. Malizia, and C. Esposito. Comparison of different commercially available cationic liposome-DNA lipoplexes: Parameters influencing toxicity and transfection efficiency. *Colloids Surf B Biointerfaces*, 68(2):136–144, Feb 2009. 5.1

[149] B. K. Kim, G. B. Hwang, Y. B. Seu, J. S. Choi, K. S. Jin, and K. O. Doh. DOTAP/DOPE ratio and cell type determine transfection efficiency with DOTAP-liposomes. *Biochim. Biophys. Acta*, 1848(10 Pt A):1996–2001, Oct 2015. 5.1

[150] T. M. Allen and L. G. Cleland. Serum-induced leakage of liposome contents. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 597(2):418–426, 1980. 5.1

[151] K Crook, BJ Stevenson, M Dubouchet, and DJ Porteous. Inclusion of cholesterol in dotap transfection complexes increases the delivery of dna to cells in vitro in the presence of serum. *Gene therapy*, 5(1), 1998. 5.1

[152] S. Yamano, J. Dai, and A. M. Moursi. Comparison of transfection efficiency of nonviral gene transfer reagents. *Mol. Biotechnol.*, 46(3):287–300, Nov 2010. 5.1, 5.4

[153] P. R. Cullis and M. J. Hope. Lipid Nanoparticle Systems for Enabling Gene Therapies. *Mol. Ther.*, 25(7), Apr 2017. 5.1, 5.6

[154] O. B. Suhr, T. Coelho, Juan Buades, J. Pouget, I. Conceicao, J. Berk, H. Schmidt, M. Waddington-Cruz, J. M. Campistol, B. R. Bettencourt, A. Vaishnaw, J. Gollob, and D. Adams. Efficacy and safety of patisiran for familial amyloidotic polyneuropathy: a phase ii multi-dose study. *Orphanet J of Rare Dis*, 10(1):109, Sep 2015. 5.1

[155] P. R. Cullis, L. D. Mayer, M. B. Bally, T. D. Madden, and M. J. Hope. Generating and loading of liposomal systems for drug-delivery applications. *Adv. Drug Delivery Rev.*, 3(3):267–282, 1989. 5.1

[156] K. J. Kauffman, J. R. Dorkin, J. H. Yang, M. W. Heartlein, F. DeRosa, F. F. Mir, O. S. Fenton, and D. G. Anderson. Optimization of Lipid Nanoparticle Formulations for mRNA Delivery in Vivo with Fractional Factorial and Definitive Screening Designs. *Nano Lett.*, 15(11):7300–7306, Nov 2015. 5.1

[157] H. Lv, S. Zhang, B. Wang, S. Cui, and J. Yan. Toxicity of cationic lipids and cationic polymers in gene delivery. *J Control Release*, 114(1):100–109, Aug 2006. 5.1

[158] N. M. Belliveau, J. Huft, P. J. Lin, S. Chen, A. K. Leung, T. J. Leaver, A. W. Wild, J. B. Lee, R. J. Taylor, Y. K. Tam, C. L. Hansen, and P. R. Cullis. Microfluidic Synthesis of Highly Potent Limit-size Lipid Nanoparticles for In Vivo Delivery of siRNA. *Mol Ther Nucleic Acids*, 1:e37, Aug 2012. 5.1

[159] I. V. Zhigaltsev, N. Belliveau, I. Hafez, A. K. Leung, J. Huft, C. Hansen, and P. R. Cullis. Bottom-up design and synthesis of limit size lipid nanoparticle systems with aqueous and triglyceride cores using millisecond microfluidic mixing. *Langmuir*, 28(7):3633–3640, Feb 2012. 5.1

[160] J. Huotari and A. Helenius. Endosome maturation. *EMBO J.*, 30(17):3481–3500, Aug 2011. 5.1

[161] T. F. Martens, K. Remaut, J. Demeester, S. C. De Smedt, and K. Braeckmans. Intracellular delivery of nanomaterials: How to catch endosomal escape in the act. *Nano Today*, 9(3):344 – 364, 2014. 5.1, 7.2, 7.2

[162] I. M. Hafez, N. Maurer, and P. R. Cullis. On the mechanism whereby cationic lipids promote intracellular delivery of polynucleic acids. *Gene Ther.*, 8(15):1188–1196, Aug 2001. 5.1

[163] M. Jayaraman, S. M. Ansell, B. L. Mui, Y. K. Tam, J. Chen, X. Du, D. Butler, L. Eltepu, S. Matsuda, J. K. Narayanannair, K. G. Rajeev, I. M. Hafez, A. Akinc, M. A. Maier, M. A. Tracy, P. R. Cullis, T. D. Madden, M. Manoharan, and M. J. Hope. Maximizing the potency of siRNA lipid nanoparticles for hepatic gene silencing in vivo. *Angew. Chem. Int. Ed. Engl.*, 51(34):8529–8533, Aug 2012. 5.1

[164] K. A. Whitehead, J. R. Dorkin, A. J. Vegas, P. H. Chang, O. Veiseh, J. Matthews, O. S. Fenton, Y. Zhang, K. T. Olejnik, V. Yesilyurt, D. Chen, S. Barros, B. Klebanov, T. Novobrantseva, R. Langer, and D. G. Anderson. Degradable lipid nanoparticles with predictable in vivo siRNA delivery activity. *Nat Commun*, 5:4277, Jun 2014. 5.1

[165] A. El-Sayed and H. Harashima. Endocytosis of gene delivery vectors: From clathrin-dependent to lipid raft-mediated endocytosis. *Mol. Ther.*, 21(6):1118–1130, Apr 2013. 5.1

[166] L. J. Pike. Rafts defined: a report on the Keystone Symposium on Lipid Rafts and Cell Function. *J. Lipid Res.*, 47(7):1597–1598, Jul 2006. 5.1

[167] J. A. Sniegowski, J. W. Lappe, H. N. Patel, H. A. Huffman, and R. M. Wachter. Base catalysis of chromophore formation in Arg96 and Glu222 variants of green fluorescent protein. *J. Biol. Chem.*, 280(28):26248–26255, Jul 2005. 5.3.1

[168] J. H. Leveau and S. E. Lindow. Predictive and interpretive simulation of green fluorescent protein expression in reporter bacteria. *J. Bacteriol.*, 183(23):6752–6762, Dec 2001. 5.3.1

[169] S. Aminikhanghahi and D. J. Cook. A Survey of Methods for Time Series Change Point Detection. *Knowl Inf Syst*, 51(2):339–367, May 2017. 5.3.2

[170] E. Uchida, H. Mizuguchi, A. Ishii-Watabe, and T. Hayakawa. Comparison of the efficiency and safety of non-viral vector-mediated gene transfer into a wide range of human cells. *Biol. Pharm. Bull.*, 25(7):891–897, Jul 2002. 5.4

[171] S. C. Semple, A. Akinc, J. Chen, A. P. Sandhu, B. L. Mui, C. K. Cho, D. W. Sah, D. Stebbing, E. J. Crosley, E. Yaworski, I. M. Hafez, J. R. Dorkin, J. Qin, K. Lam, K. G. Rajeev, K. F. Wong, L. B. Jeffs, L. Nechev, M. L. Eisenhardt, M. Jayaraman, M. Kazem, M. A. Maier, M. Srinivasulu, M. J. Weinstein, Q. Chen, R. Alvarez, S. A. Barros, S. De, S. K. Klimuk, T. Borland, V. Kosovrasti, W. L. Cantley, Y. K. Tam, M. Manoharan, M. A. Ciufolini, M. A. Tracy, A. de Fougerolles, I. MacLachlan, P. R. Cullis, T. D. Madden, and M. J. Hope. Rational design of cationic lipids for siRNA delivery. *Nat. Biotechnol.*, 28(2):172–176, Feb 2010. 5.5

[172] Y. Y. Tam, S. Chen, and P. R. Cullis. Advances in Lipid Nanoparticles for siRNA Delivery. *Pharmaceutics*, 5(3):498–507, Sep 2013. 5.5

[173] A. Zidovska, H. M. Evans, A. Ahmad, K. K. Ewert, and C. R. Safinya. The role of cholesterol and structurally related molecules in enhancing transfection of cationic liposome-DNA complexes. *J Phys Chem B*, 113(15):5208–5216, Apr 2009. 5.6

[174] H. Yasar, A. Biehl, C. De Rossi, M. Koch, X. Murgia, B. Loretz, and C. M. Lehr. Kinetics of mRNA delivery and protein translation in dendritic cells using lipid-coated PLGA nanoparticles. *J Nanobiotechnology*, 16(1):72, Sep 2018. 5.6, 5.6

[175] M. P. Monopoli, C. Aberg, A. Salvati, and K. A. Dawson. Biomolecular coronas provide the biological identity of nanosized materials. *Nat Nanotechnol*, 7(12):779–786, Dec 2012. 5.6

[176] F. Wang, L. Yu, M. P. Monopoli, P. Sandin, E. Mahon, A. Salvati, and K. A. Dawson. The biomolecular corona is retained during nanoparticle uptake and protects the cells from the damage induced by cationic nanoparticles until degraded in the lysosomes. *Nanomedicine*, 9(8):1159–1168, Nov 2013. 5.6

[177] S. Motta, V. Rondelli, L. Cantu, E. Del Favero, M. Aureli, D. Pozzi, G. Caracciolo, and P. Brocca. What the cell surface does not see: The gene vector under the protein corona. *Colloids Surf B Biointerfaces*, 141:170–178, May 2016. 5.6

[178] G Caracciolo, D Pozzi, AL Capriotti, C Cavaliere, S Piovesana, G La Barbera, A Amici, and A Laganà. The liposome–protein corona in mice and humans and its implications for in vivo delivery. *Journal of Materials Chemistry B*, 2(42):7419–7428, 2014. 5.6

[179] U. Sakulkhu, M. Mahmoudi, L. Maurizi, G. Coullerez, M. Hofmann-Amtenbrink, M. Vries, M. Motazacker, F. Rezaee, and H. Hofmann. Significance of surface charge and shell material of superparamagnetic iron oxide nanoparticle (SPION) based core/shell nanoparticles on the composition of the protein corona. *Biomater Sci*, 3(2):265–278, Feb 2015. 5.6

[180] A. Akinc, W. Querbes, S. De, J. Qin, M. Frank-Kamenetsky, K. N. Jayaprakash, M. Jayaraman, K. G. Rajeev, W. L. Cantley, J. R. Dorkin, J. S. Butler, L. Qin, T. Racie, A. Sprague, E. Fava, A. Zeigerer, M. J. Hope, M. Zerial, D. W. Sah, K. Fitzgerald, M. A. Tracy, M. Manoharan, V. Koteliansky, A. Fougerolles, and M. A. Maier. Targeted Delivery of RNAi Therapeutics With Endogenous and Exogenous Ligand-Based Mechanisms. *Mol. Ther.*, 18(7):1357–1364, Jul 2010. 5.6

[181] T. Hishiki, Y. Shimizu, R. Tobita, K. Sugiyama, K. Ogawa, K. Funami, Y. Ohsaki, T. Fujimoto, H. Takaku, T. Wakita, T. F. Baumert, Y. Miyanari, and K. Shimotohno. Infectivity of hepatitis C virus is influenced by association with apolipoprotein E isoforms. *J. Virol.*, 84(22):12048–12057, Nov 2010. 5.6

[182] A. K. Leung, I. M. Hafez, S. Baoukina, N. M. Belliveau, I. V. Zhigaltsev, E. Afshinmanesh, D. P. Tieleman, C. L. Hansen, M. J. Hope, and P. R. Cullis. Lipid Nanoparticles Containing siRNA Synthesized by Microfluidic Mixing Exhibit an Electron-Dense Nanostructured Core. *J Phys Chem C Nanomater Interfaces*, 116(34):18440–18450, Aug 2012. 5.6, 7.3.1

[183] K. Frank. Small-angle X-ray scattering analysis of DNA origami lattices and lipid bulk phases. Master's thesis, Ludwig-Maximilinas-University Munich, 2017. 5.6

[184] A. Murschhauser, P. J. F. Röttgermann, D. Woschée, M. F. Ober, Y. Yan, K. A. Dawson, and J. O. Rädler. A high-throughput microscopy method for single-cell analysis of event-time correlations in nanoparticle-induced cell death. *Communications biology*, 2(1):35, 2019. 5.6, 7, 7.2, 7.2

[185] B. Li, X. Luo, and Y. Dong. Effects of Chemically Modified Messenger RNA on Protein Expression. *Bioconjug. Chem.*, 27(3):849–853, Mar 2016. 6

[186] C. J. Wilusz, M. Wormington, and S. W. Peltz. The cap-to-tail guide to mRNA turnover. *Nat. Rev. Mol. Cell Biol.*, 2(4):237–246, Apr 2001. 6, 6.4

[187] S. Holtkamp, S. Kreiter, A. Selmi, P. Simon, M. Koslowski, C. Huber, O. Tureci, and U. Sahin. Modification of antigen-encoding RNA increases stability, translational efficacy, and T-cell stimulatory capacity of dendritic cells. *Blood*, 108(13):4009–4017, Dec 2006. 6

[188] A. E. Grier, S. Burleigh, J. Sahni, C. A. Clough, V. Cardot, D. C. Choe, M. C. Krutein, D. J. Rawlings, M. C. Jensen, A. M. Scharenberg, and K. Jacoby. pEVL: A Linear Plasmid for Generating mRNA IVT Templates With Extended Encoded Poly(A) Sequences. *Mol Ther Nucleic Acids*, 5:e306, Apr 2016. 6

[189] T. Wada and A. Becskei. Impact of Methods on the Measurement of mRNA Turnover. *Int J Mol Sci*, 18(12), Dec 2017. 6

[190] C. Y. Chen and A. B. Shyu. Mechanisms of deadenylation-dependent decay. *Wiley Interdiscip Rev RNA*, 2(2):167–183, 2011. 6.1.1, 6.1.2, 6.4

[191] D. R. Schoenberg and L. E. Maquat. Regulation of cytoplasmic mRNA decay. *Nat. Rev. Genet.*, 13(4):246–259, Mar 2012. 6.1.1

[192] D. Cao and R. Parker. Computational modeling of eukaryotic mrna turnover. *Rna*, 7(9):1192–1212, 2001. 6.1.1

[193] J. Peng, E. L. Murray, and D. R. Schoenberg. In vivo and in vitro analysis of poly(A) length effects on mRNA translation. *Methods Mol. Biol.*, 419:215–230, 2008. 6.1.2

[194] D. A. Mangus, M. C. Evans, and A. Jacobson. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.*, 4(7):223, 2003. 6.1.2

[195] L. Weill, E. Belloc, F. A. Bava, and R. Mendez. Translational control by changes in poly(A) tail length: recycling mRNAs. *Nat. Struct. Mol. Biol.*, 19(6):577–585, Jun 2012. 6.1.2

[196] H. Chang, J. Lim, M. Ha, and V. N. Kim. TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. *Mol. Cell*, 53(6):1044–1052, Mar 2014. 6.1.2, 6.4

[197] A. O. Subtelny, S. W. Eichhorn, G. R. Chen, H. Sive, and D. P. Bartel. Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature*, 508(7494):66–71, Apr 2014. 6.1.2, 6.3.2, 6.4

[198] Z. Trepotec, J. Geiger, C. Plank, M. K. Aneja, and C. Rudolph. Segmented poly(A)

tails significantly reduce recombination of plasmid DNA without affecting mRNA translation efficiency or half-life. *RNA*, 25(4):507–518, Apr 2019. 6.1.3, 6.3.2

[199] D. Pumpernik and B. Oblak, B.and Borštnik. Replication slippage versus point mutation rates in short tandem repeats of the human genome. *Molecular Genetics and Genomics*, 279(1):53–61, 2008. 6.1.3

[200] M. Beverly, C. Hagen, and O. Slack. Poly A tail length analysis of in vitro transcribed mRNA by LC-MS. *Anal Bioanal Chem*, 410(6):1667–1677, Feb 2018. 6.1.3, 6.3.2

[201] L. A. Castellano and A. A. Bazzini. Poly (a) tails: longer is not always better. *Nat. Struct. Mol. Biol.*, 24(12):1010, 2017. 6.3.2

[202] T. Preiss, M. Muckenthler, and M. W. Hentze. Poly (a)-tail-promoted translation in yeast: implications for translational control. *Rna*, 4(11):1321–1331, 1998. 6.4

[203] S. A. Lima, L. B. Chipman, A. L. Nicholson, Y. H. Chen, B. A. Yee, G. W. Yeo, J. Coller, and A. E. Pasquinelli. Short poly(A) tails are a conserved feature of highly expressed genes. *Nat. Struct. Mol. Biol.*, 24(12):1057–1063, Dec 2017. 6.4

[204] J. Lim, M. Ha, H. Chang, S. C. Kwon, D. K. Simanshu, D. J. Patel, and V. N. Kim. Uridylation by TUT4 and TUT7 marks mRNA for degradation. *Cell*, 159(6):1365–1376, Dec 2014. 6.4

[205] N. Xu, C. Y. Chen, and A. B. Shyu. Modulation of the fate of cytoplasmic mRNA by AU-rich elements: key sequence features controlling mRNA deadenylation and decay. *Mol. Cell. Biol.*, 17(8):4611–4621, Aug 1997. 6.4

[206] D. Zheng, N. Ezzeddine, C.-Y. A. Chen, W. Zhu, X. He, and A.-B. Shyu. Deadenylation is prerequisite for p-body formation and mrna decay in mammalian cells. *J Cell Biol*, 182(1):89–101, 2008. 6.4

[207] H. K. Karnati, R. S. Yalagala, R. Undi, S. R. Pasupuleti, and R. K. Gutti. Therapeutic potential of siRNA and DNAzymes in cancer. *Tumour Biol.*, 35(10):9505–9521, Oct 2014. 7.1

[208] V. Hornung, F. Bauernfeind, A. Halle, E. O. Samstad, H. Kono, K. L. Rock, K. A. Fitzgerald, and E. Latz. Silica crystals and aluminum salts activate the NALP3 inflammasome through phagosomal destabilization. *Nat. Immunol.*, 9(8):847–856, Aug 2008. 7.2, 7.2

[209] S. T. Yang, E. Zaitseva, L. V. Chernomordik, and K. Melikov. Cell-penetrating peptide induces leaky fusion of liposomes containing late endosome-specific anionic lipid. *Biophys. J.*, 99(8):2525–2533, Oct 2010. 7.2, 7.2

[210] Z. ur Rehman, D. Hoekstra, and I. S. Zuhorn. Mechanism of polyplex- and lipoplex-mediated delivery of nucleic acids: real-time visualization of transient membrane destabilization without endosomal lysis. *ACS Nano*, 7(5):3767–3777, May 2013. 7.2

[211] Thermo Fisher. *pHodro Red and Green Dextran for Endocytosis*, Aug 2012. 7.2

[212] Thermo Fisher. *LysoTracker and LysoSensor Probes*, 2.0 edition. 7.2

[213] D. K. Bonner, C. Leung, J. Chen-Liang, L. Chingozha, R. Langer, and P. T. Hammond. Intracellular trafficking of polyamidoamine-poly(ethylene glycol) block

copolymers in DNA delivery. *Bioconjug. Chem.*, 22(8):1519–1525, Aug 2011. 7.2, 7.4

[214] L. Anhäuser, S. Huwel, T. Zobel, and A. Rentmeister. Multiple covalent fluorescence labeling of eukaryotic mRNA at the poly(A) tail enhances translation and can be performed in living cells. *Nucleic Acids Res.*, Feb 2019. 7.3

[215] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox. Real-time kinetics of gene activity in individual bacteria. *Cell*, 123(6):1025–1036, Dec 2005. 7.3

[216] Y. Ben-Ari, Y. Brody, N. Kinor, A. Mor, T. Tsukamoto, D. L. Spector, R. H. Singer, and Y. Shav-Tal. The life of an mRNA in space and time. *J. Cell. Sci.*, 123(Pt 10):1761–1774, May 2010. 7.3

[217] B. Wu, A. R. Buxbaum, Z. B. Katz, Y. J. Yoon, and R. H. Singer. Quantifying Protein-mRNA Interactions in Single Live Cells. *Cell*, 162(1):211–220, Jul 2015. 7.3

[218] H. E. Johansson, L. Liljas, and O. C. Uhlenbeck. Rna recognition by the ms2 phage coat protein. In *Seminars in Virology*, volume 8, pages 176–185. Elsevier, 1997. 7.3

[219] S. Tyagi and F. R. Kramer. Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.*, 14(3):303–308, Mar 1996. 7.3

[220] P. J. Santangelo, B. Nix, A. Tsourkas, and G. Bao. Dual FRET molecular beacons for mRNA detection in living cells. *Nucleic Acids Res.*, 32(6):e57, Apr 2004. 7.3

[221] M. You and S. R. Jaffrey. Structure and Mechanism of RNA Mimics of Green Fluorescent Protein. *Annu Rev Biophys*, 44:187–206, 2015. 7.3

[222] J. Zhang, J. Fei, B. J. Leslie, K. Y. Han, T. E. Kuhlman, and T. Ha. Tandem Spinach Array for mRNA Imaging in Living Bacterial Cells. *Sci Rep*, 5:17295, Nov 2015. 7.3

[223] E. V. Dolgosheina, S. C. Jeng, S. S. Panchapakesan, R. Cojocaru, P. S. Chen, P. D. Wilson, N. Hawkins, P. A. Wiggins, and P. J. Unrau. RNA mango aptamer-fluorophore: a bright, high-affinity complex for RNA labeling and tracking. *ACS Chem. Biol.*, 9(10):2412–2420, Oct 2014. 7.3

[224] A. Raj, P. van den Bogaard, S. A. Rifkin, A. van Oudenaarden, and S. Tyagi. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods*, 5(10):877–879, Oct 2008. 7.3.2, 7.3.2

[225] S. O. Skinner, L. A. Sepulveda, H. Xu, and I. Golding. Measuring mRNA copy number in individual Escherichia coli cells using single-molecule fluorescent in situ hybridization. *Nat Protoc*, 8(6):1100–1113, Jun 2013. 7.3.2

[226] H. Xu, L. A. Sepulveda, L. Figard, A. M. Sokac, and I. Golding. Combining protein and mRNA quantification to decipher transcriptional regulation. *Nat. Methods*, 12(8):739–742, Aug 2015. 7.3.2

[227] J. M. Levsky and R. H. Singer. Fluorescence in situ hybridization: past, present and future. *J. Cell. Sci.*, 116(Pt 14):2833–2838, Jul 2003. 7.3.2

[228] F. Mueller, A. Senecal, K. Tantale, H. Marie-Nelly, N. Ly, O. Collin, E. Basyuk, E. Bertrand, X. Darzacq, and C. Zimmer. FISH-quant: automatic counting of transcripts in 3D FISH images. *Nat. Methods*, 10(4):277–278, Apr 2013. 7.3.2

[229] A. L. Szymczak, C. J. Workman, Y. Wang, K. M. Vignali, S. Dilioglou, E. F. Vanin, and D. A. Vignali. Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat. Biotechnol.*, 22(5):589–594, May 2004. 7.4

[230] J. H. Kim, S. R. Lee, L. H. Li, H. J. Park, J. H. Park, K. Y. Lee, M. K. Kim, B. A. Shin, and S. Y. Choi. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS ONE*, 6(4):e18556, 2011. 7.4

[231] E. K. Larsen, M. B. Mikkelsen, and N. B. Larsen. Protein and cell patterning in closed polymer channels by photoimmobilizing proteins on photografted poly(ethylene glycol) diacrylate. *Biomicrofluidics*, 8(6):064127, Nov 2014. 8

[232] T. Falk, D. Mai, R. Bensch, O. Cicek, A. Abdulkadir, Y. Marrakchi, A. Bohm, J. Deubner, Z. Jackel, K. Seiwald, A. Dovzhenko, O. Tietz, C. Dal Bosco, S. Walsh, D. Saltukoglu, T. L. Tay, M. Prinz, K. Palme, M. Simons, I. Diester, T. Brox, and O. Ronneberger. U-Net: deep learning for cell counting, detection, and morphometry. *Nat. Methods*, 16(1):67–70, 2019. 8

[233] Biosearch Technologies. *Stellaris RNA FISH Protocol for adherent cells*, Apr 2015. A

Danksagung

An dieser Stelle ein herzlicher Dank an alle, die in irgendeiner Form mir bei der Entstehung dieser Arbeit geholfen haben und mich während der Promotion unterstützt haben. Besonders bedanken möchte ich mich hierfür bei:

Joachim Rädler für die Betreuung dieser Arbeit und die viele Unterstützung in den letzten vier Jahren. Ich habe viel gelernt!

QBM für die Unterstützung durch das Stipendium, die Retreats (besonders Venedig), die Finanzierung zur Teilnahme an Konferenzen und Forschungsaufenthalten.

Meinen Kooperationspartnern für die gute Zusammenarbeit: Fabian und Jan für die Arbeit am "Modeling" Paper; Helmut für deinen wertvollen Beitrag zum "Onset" Paper und auch für den interessanten Besuch in Stony Brook; Aleksandra und Lennart für die spannende Arbeit mit den LNPs; Mehrije, Zeljka, Manish, Carsten und Christian für die Bereitstellung vieler mRNA Konstrukte und die Arbeit am "poly(A)" Projekt.

Den Studenten, die ich betreuen durfte, deren wissenschaftlicher Beitrag mit in dieser Arbeit steckt: Sophia G. für die wertvolle Untersuchung verschiedener Mikrostrukturierungen, Neha für die Untersuchung verschiedener Transfektionsmittel, Tobias für die Programmierung des "Clustering approachs", Joachim für die Hilfe bei der Etablierung der FISH Methode und Simone für die großartige Arbeit am poly(A) Projekt.

Dem gesamten Lehrstuhl Rädler und besonders der Arbeitsgruppe Rädler mit all seinen zahlreichen Mitgliedern über die letzten Jahre. Sonja, vielen lieben Dank für die ganzen Pilatesstunden. Janina, für ein bisschen Erlangen in München. Daniel, ein großes Dankeschön für deine Unterstützung bei der Datenanalyse. Besonderer Dank an Gerlinde und Charlott für die Hilfe aller Art. Ein herzlicher Dank an alle, die als Bürokollegen die Promotionszeit unvergesslich gemacht haben: Tobi, Judith, Matthias, Bene, Sophia, Christoph, Miriam, Alexandra und Merlin und ganz besonders dem harten Kern!

Meinen fleißigen Korrekturlesern: Sophia G., Simone, Miriam, Matthias, Christoph, Alexandra, Sophia S. und ganz besonders Judith.

Allen Freunden die dafür gesorgt haben, dass die Freizeit nicht zu kurz kommt und sich meine Erzählungen von Zellen schubsen und zum Leuchten bringen angehört haben.

Gerda und Dieter für ihre Unterstützung.

Meiner Familie auf die ich mich immer verlassen kann. Ganz besonders meinen Eltern ohne die ich meine Unilaufbahn nicht geschafft hätte. Liebe Mutti, vielen lieben Dank, dass du immer alles möglich machst was möglich ist und noch mehr.

Basti, ohne deinen Rückhalt wäre die Arbeit eventuell nie entstanden. Vielen Dank, einfach für alles!